Shoulder complaints

The occurrence, course and diagnosis

Jolanda Luime

This thesis was printed with support of the Erasmus University Medical Centre, Netherlands Expert Centre for Work -Related Musculoskeletal Disorders, and the Anna Fonds.

Cover design: Ellen Spanjaard, Bilthoven
Cover photography: Judith Kuiper, Amsterdam
Printing: Optima BV, Rotterdam

ISBN: 90-8559-007-8

Shoulder complaints

The occurrence, course and diagnosis

Schouderklachten

Frequentie, beloop en diagnose

Proefschrift

Ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus

Prof. Dr. S.W.J. Lamberts

en volgens besluit van het College van Promoties.
De openbare verdediging zal plaatsvinden op

donderdag 16-12-2004 om 16.00 uur

door Johanna Jacoba Luime geboren te Ouddorp

Promotie commissie

Promotoren:

Prof. dr. Bart W. Koes

Prof. dr. Jan A.N. Verhaar

Overige leden: Prof dr. J. Mieke H. Hazes

Prof. dr. Myriam G.M. Hunink

Prof. dr. Henk J. Stam

Copromotor:

Dr. ir. Alex Burdorf

Contents

Chapter 1

General introduction

1 Introduction

Shoulder complaints are expressed in a variety of symptoms.[1] In many cases, the prominent symptom is pain. In some cases, pain is present most of the day and frequently also at night.[2,3] In other cases, it is provoked primarily by physical activities. Often it is accompanied by restricted range in shoulder movement. This inhibit people in their daily lives by reducing their ability to use their arm in activities such as dressing, personal hygiene, work, household activities, hobbies and sports.[2] These shoulder symptoms may also lead to sick leave, bringing costs to the workers themselves, as well as to employers and society.[4] Although these are a common musculoskeletal complaint, the estimated prevalence of shoulder complaints varies considerably. Reported prevalences in the general population differ from 6% up to 25% [5,6] and little is known about the incidence in this population.[5] The wide range in reported prevalence hampers a considered estimation of the true problem of shoulder complaints, and thereby complicates the assessment of the proportion of different subgroups in the spectrum of shoulder complaints in the general population (i.e. the proportion which seeks care, uses medication or is absent from work).

The extent of shoulder problems is determined not only by the level of their occurrence, but also by their course. In highly prevalent but acute self-limiting disorders, a wait-and-see policy is presumably enough to manage the complaint. However, when the course of shoulder complaints is characterised by persistence or frequently recurrent episodes, this may not be sufficient. The complaints suffered by patients visiting the general practitioner seem to be neither short lived nor characterised by isolated episodes. Half to 70% of the patients in primary care settings reported the persistence of a complaint after 6 months and 40% to 50% after 1 year.[2,7,8] Twenty to fifty percent of these patients had had earlier episodes of shoulder complaints.[2,7] Persistence of shoulder complaints was also found in a working population in the forest industry.[9]

Data on the course of shoulder complaints in more general populations is missing. To understand shoulder problems, more information is needed both on their course over time and on the characteristics of recurrent complaints. Since not all subjects with shoulder complaints undergo recurrent or persistent complaints [2, 7-9] greater insight is required into subgroups undergoing a higher risk of recurrent shoulder complaints. And although previous studies on shoulder pain have focused on the risk factors for the onset of complaints, it is uncertain whether their information can be extrapolated to recurrent events.

The underlying pathomechanisms for most specific shoulder disorders are poorly understood. Because the glenohumeral joint, the acromioclavicular joint, the sternoclavicular joint, the scapula, the surrounding ligaments, muscles, nerves and blood vessels contain nociceptive fibers, it is not unfair to state that all structures of the shoulder can be a source of pain.[10] However, the originating mechanisms and specific symptoms related to the different structures of the shoulder remain unclear.

Although knowledge of its pathology is relatively limited, approaches in clinical practice are often based on searching for underlying pathological substrates. History-taking and clinical examination are commonly used as diagnostic instruments to differentiate between disorders of the muscle tendons, capsule, ligaments or other structures of the shoulder. However, previous studies on the reliability of clinical test have shown that agreement on the presence of symptoms and underlying sources of pain is poor to moderate among physiotherapists, and medical specialists.[11,12] And moreover, there is no clear overview of the validity of history items and clinical tests related to specific shoulder structures.

Thus, in summary, despite the impact of shoulder complaints on patients, employers and society, understanding of its occurrence is inconsistent. Neither is much known about the course of these complaints over time, or about the validity of history-taking and physical examination for specific shoulder disorders.

2 Objectives of this thesis

The primary objectives of this thesis are:

- To describe the incidence and prevalence of shoulder pain in the general population.

- To explore the course of shoulder complaints and to assess the influence of risk factors on the incidence and recurrence of these complaints.

- To evaluate the accuracy and informativeness of history-taking and clinical examination for diagnosing specific shoulder disorders.

3 Outline of the thesis

Following this general introduction to the background and objectives of this study, Chapter 2 describes and discusses the epidemiological evidence on the prevalence and incidence of shoulder complaints by presenting a systematic review of epidemiological studies on shoulder pain. When evaluating differences between estimates of the prevalence of shoulder complaints, the methodological quality of the published studies is also considered, and the influence of different case definitions on the estimates is discussed.
Chapter 3 presents the results of a two-year follow-up study on the prevalence, incidence, and recurrence of neck and shoulder pain in a working population. Data of a longitudinal study on musculoskeletal disorders are used to explore the relationships between incidence, recurrence, prevalence and potential risk factors.
Chapter 4 continues the study on the course of neck and shoulder complaints by asking whether the risk factors for recurrent episodes of such complaints are the same as for incident episodes.

In Chapters 5 and 6, the focus shifts to the diagnosis of shoulder pain. Because the cornerstones of diagnostic management in health care are the history of the complaint and clinical findings in the physical examination, we present a systematic review of the validity of history-taking and clinical tests for shoulder pain. Chapter 5 deals with diagnosis instability and intra-articular pathology (e.g. labral tears). Chapter 6 examines the validity of tests for diagnosing rotator cuff tears and impingement.

Finally, Chapter 7 presents the main conclusions of the previous chapters within the light of the three objectives of this thesis. This chapter pays attention to the limitations of our study, the methodological consequences of the episodic course of shoulder complaints, the influence of the case definition on outcome, and the influence of diagnostic measures in the management of shoulder pain.

4 References

1. van der Heijden GJ. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
2. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. BMJ 1996;313:601-602.
3. Van der Windt DA, Koes BW, de Jong BA. Shoulder disorders in general practice: incidence, patient characteristics, and management. Ann Rheum Dis 1995; 54:959-964.
4. Silverstein B, Welp E, Nelson N, Kalat J. Claims incidence of work-related disorders of the upper extremities: Washington State, 1987 through 1995. Am J Pub Health 1998;88:1827-1833.
5. Bjelle A. Epidemiology of shoulder problems. Baillieres Clin Rheumatol 1989;3:437-451.
6. Badley EM, Tennant A. Changing profile of joint disorders with age: findings from a postal survey of the population of Calderdale, West Yorkshire, United Kingdom. Ann Rheum Dis 1992;51:366-371.
7. Van der Windt DA, Koes BW, Boeke AJ, Deville W, de Jong BA, Bouter Lm. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-523.

8. Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-de Jong B. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology (Oxford) 1999;38:160-163.
9. Miranda H, Viikari-Juntura E, Martikainen R, Takala EP, Riihimaki H. A prospective study of work related factors and physical exercise as predictors of shoulder pain. Occup Environ Med 2001;58:528-534.
10. Vangsness CT, Ennis M, Taylor JG, Atkinson R. Neural anatomy of the glenohumeral ligaments, labrum and subacromial bursa. Arthroscopy 1995;11:180-184.
11. de Winter AF, Jans MP, Scholten RJ, Deville W, van Schaardenburg D, Bouter LM. Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. Ann Rheum Dis 1999;58:272-277.
12. Norregaard J, Krogsgaard MR, Lorenzen T, Jensen EM. Diagnosing patients with longstanding shoulder joint pain. Ann Rheum Dis 2002;61:646-649.

Chapter 2

Prevalence and incidence of shoulder pain in the general population; a systematic review

Published as: Luime JJ, Koes BW, Hendriksen IJ, Burdorf A, Verhagen AP, Miedema HS, Verhaar JA. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol 2004;33:73-81

Abstract

Objective To investigate the incidence and prevalence of shoulder complaints in the general population. **Method** A systematic review of the literature was conducted. MEDLINE, EMBASE, and CINAHL were searched for relevant studies. **Results** Eighteen studies on prevalence and one study on incidence met the inclusion criteria. Incidence figures of 0.9 - 2.5% were found for different age groups. Prevalence figures differed from 6.9 - 26% for point prevalence, 18.6 - 31% for one-month prevalence, 4.7 - 46.7% for one-year prevalence and 6.7 - 66.7% for life-time prevalence. Prevalence rates decreased when the case definition was restricted in terms of duration of pain or the presence of limited movements and increased when the location for pain was enlarged. **Conclusion** The reported prevalence figures on shoulder complaints diverged strongly. Health professionals and policymakers who estimate the amount of medical care needed and related cost should be aware of the variations in prevalence rate and the underlying reasons for these differences.

1 Introduction

Shoulder pain is an important medical and socio-economic problem in western society. Pain and stiffness in the shoulder may lead to inability to work and/or to carry out household and leisure time activities, burdening both patient and society. For many patients, shoulder complaints are not self-limiting within weeks or months; about 50% of the patients, who visited a general practitioner, still reported complaints after 12 months.[1]

The literature reports large ranges of incidence and prevalence rates.[1,2] Difference in case definitions, types of sampling procedures, variety in response rates and type of instruments used may be responsible for these large ranges. Strong diversity in reported prevalence and incidence rates may cause problems in the estimation of the magnitude of the problem, as well as estimation of the amount of medical care needed and their related costs.

To gain insight in the large ranges of incidence and prevalence figures we reviewed the literature systematically for studies on shoulder pain in the general population.

2 Methods

Study selection MEDLINE (1966-2001), EMBASE (1980-2001) and CINAHL (1982-2001) were searched for identification of relevant studies. The search is based on a modified strategy used by Green et al.[3] (keywords: shoulder (exploded), glenohumeral, scapula, clavicula, acromion, rotator cuff, supraspinatus, supra-spinatus, infraspinatus, infra-spinatus, serratus anterior, subscapularis, not cancer, not animal[mesh], prevalence, incidence). The search had no language restrictions. In addition, the references in relevant publications were also examined.

Selection of studies was based on two-stage process. Firstly, all abstracts or titles found by the electronic searches were scrutinized by JJL. Secondly, after obtaining copies of eligible papers, IH and JJL independently assessed all these articles for inclusion and exclusion criteria. Minimal requirements for inclusion were shoulder complaints, cross sectional study design for prevalence studies and longitudinal study design for incidence studies, data of the general population of 18 years and older, and in the incidence studies the subjects had to be symptom-free. Studies were excluded if (I) the population suffered from specific underlying pathology such as tumours, trauma (fractures), infection, inflammatory disorders (rheumatoid arthritis), etc and (II) if they were published in non-scientific journals (such as reports based on governmental databases). These reports are not systematically indexed in a database and therefore difficult to obtain. To avoid selection bias we excluded these reports from the review.

Assessment of methodological quality IH, JL and LB independently assessed the methods of data collection and evaluated the response rate by a four item quality list, which is described in table 1. The items were based on generally accepted principles for observational studies. If one or more items were scored negative the study results are possibly biased. The outcome of these studies should be interpreted with caution. If there were disagreements between the reviewers on the quality assessment, these were resolved by consensus.

Table 1
Description of the criteria for assessment of the methodological quality of the incidence and prevalence studies

Criterion	Description
Random sampling	The sample was taken randomly from the population or the whole study population was approached. The method of sampling was described in the article or there was a reference in the text.

Criterion	Description
Operational criteria for establishing complaints or disorder	The complaint, disorder or diagnosis was determined by predefined criteria. These criteria had to be specified in a way that they are reproducible by others.
Use of valid and reliable measurements	There was a reference to a validation and reliability study, or the method of validation is described in the article. If references were available these were checked by JL for being truly a validation or reliability study.
Response rate	The response rate was 60% or higher.

Outcome of the studies and statistical pooling Outcome of studies on prevalence and incidence included simple frequency enumeration of shoulder complaints. For each study details were extracted on study population (setting, sampling, response rate) and outcome (case definition, prevalence or incidence). Pooled incidence or prevalence was calculated only if there was homogeneity across studies considering instruments, case definitions, age groups and spells.

3 Results

We identified 1461 citations for the prevalence and 1688 for the incidence of shoulder pain from the electronic search, and obtained full papers for 42 of them. The reference lists of these studies revealed 33 additional studies. A total of 17 studies[4,6-21] met the inclusion criteria for prevalence studies and one study [5] met the inclusion criteria for both prevalence and incidence studies.
Fifty eight studies were excluded: 30 studies presented no prevalence or incidence numbers for shoulder complains, 11 provided data on the combination of neck and shoulder complaints, 8 were not based on the general population, 4 comprehended systemic disorders, 4 used data of studies which were already included and one study was a review.
The 18 studies were assessed for their methodological quality (Table 2) by IH, JL and LB. Disagreement occurred on 32 items (44%), mostly

related to the item for establishing the disorder. In a consensus meeting the reviewers came to agreement on all items. The item random sampling was fulfilled by 15[4-12,14,16-19] out of the 18 studies. Fifteen studies[4-13,15,17-19,21] defined their criteria for establishing shoulder complaints or disorders clearly, in two studies[14,16] the criteria were not fully specified and in one study [20] they were missing. The response rate was in thirteen studies[4-9,14-16-19,21] 60% or higher, in one study it was lower,[14] in three studies[10,11,13] it was missing and in one unclear.[12] Valid and reliable measurements were used in two studies.[6,18]

Table 2

Result of the quality assessment of the selected studies.

First author	Random sampling	Criteria clear and operational for establishing disorder	Response > 60%	Instrument valid and reliable
Andersson et al. 1993 [6]	+	+	+	+
Natvig et al. 1994 [18]	+	+	+	+
Brattberg et al. 1996 [9]	+	+	+	?
Badley et al. 1992 [7]	+	+	+	?
Mullerdorf et al. 2000 [17]	+	+	+	?
Urwin et al. 1998 [21]	+	+	+	?
Brattberg et al. 1989 [8]	+	+	+	?
Adebajo et al. 1992 [4]	+	+	+	-
Pope et al. 1997 [19]	+	+	+	-
Allander 1974 [5]	+	+	+	-
Chard et al. 1991 [10]	+	+	?	-
Makela et al. 1993 [15]	-	+	+	?
Jacobsson et al. 1989 [14]	+	?	+	-
Eriksen et al. 1998 [12]	+	+	?	?
Meyers et al. 1982 [16]	+	?	+	-
Cunningham et al. 1984 [11]	+	+	?	?
Reyes et al. 2000 [20]	-	-	+	?
Gomez et al. 1997 [13]	-	+	?	-

+ = positive score; - = negative score; ? = unclear

Studies Thirteen studies[4,5,7,9-11,13,14-16,18,20,21] presented data on the prevalence of shoulder pain, 4 studies[6,8,12,17] on shoulder/arm pain and

one on both.[19] Due to the differences in case definitions, spells and age we refrained from pooling the results, so we choose to describe the results in this section. The studies were classified in point and period prevalence. Study characteristics are presented in Table 3 and the outcome in Table 4

Prevalence of shoulder pain The point prevalence of shoulder pain among adults younger than 70 years ranged from 7 - 27% [4,5,7,13] and for the adults older than 70 from 13.2%-26% [5,7,10]. The one-month prevalence ranged from 19 - 31% [15,19,21]. The one-year prevalence ranged from 5 – 47% [9,11,14,18], although the study of Brattberg et al.[9] included only people older than 76 years (prevalence: 34,5%). Lifetime prevalence ranged from 7 - 67%.[4,11,16,20] The prevalence within a specific period of time decreased when the case definition requested not only the presence of subjective complaints, but also demanded the presence of limited motion (see Jacobson[14] versus Natvig[18]) and/or the expansion of the duration of the pain episode (see Urwin[21] versus Makela [15] and Pope [19]). The prevalence seemed to increase with age,[5,7,21] being a women,[18,21] and expansion of the anatomical area upon the case definition is based.[19]

Prevalence of shoulder-arm pain The one-month prevalence of shoulder arm pain was around 33%.[11,20] Life time prevalence was estimated (in only one study) to be 30.7%.[21] The one-year prevalence of chronic shoulder-arm pain was estimated between 8.4% and 20%.[22,23] Similar to shoulder pain, the prevalence of shoulder-arm pain seemed to increase with age and women reported more often the presence of complaints than men.

Table 3

Description of the population and sample in the selected studies.

	Study population[†]				Sample[Ψ]			
Study	Population source	country	Age (years)	N	N	type	Respons rate	% females
Badley et al. 1992 [7]	Population of Calderdale	UK	16-85+	189,402	42,829	R	87%	-
Gomez et al. 1997 [13]	A county of Spain	Spain	40-69	-	982	NR	-	45.3%
Allander 1974 [5]	Population of Stockholm,	Sweden	31-35, 42-46, 56-60, 70-74	-	4195	NR	78%	-
Adebajo et al. 1992 [4]	Population of Lagun,	Nigeria	-	-	100	R	100%	-
Chard et al. 1991 [10]	Population UK	UK	70+	-	644	R	-	50.4%
Urwin et al. 1998 [21]	Population of Tameside/ Glossop area	UK	16- 75	30,000	5752	R	78.5%	52%
Pope et al. 1997 [19]	Population of Stockport	UK	-	-	312	R	66%	55%
Makela et al. 1999 [15]	Population of Finland,	Finland	30+	-	7042	NR	90%	-
Jacobsson et al. 1989 [14]	Population of Malmo	Sweden	-	230,000	445	R	49.4%	-
Natvig et al. 1994 [18]	Population of Ullensaker	Sweden	20-22, 30-32, 40-42, 50-52, 60-62, 70-72	18,000	2726	R	67.3%	52%

Study	Study population[†] Population source	country	Age (years)	N	Sample[Ψ] N	type	Respons rate	% females
Brattberg et al. 1996[9]	Population Sweden	Sweden	76+	6524	537	-	95 %	-
Cunningham et al. 1984[11]	Population of the USA	USA	-	-	6913	R	-	-
Reyes et al. 2000 [20]	Population of Cuba	Cuba	-	11.14 million	300	NR	100%	41.3%
Meyers et al. 1982[16]	Population Rietpoort and Houtbay	South-Africa	65+	184	162	WP	88%	52%
Andersson et al. 1993[6]	Population of Bromolla and Simrishamn,	Sweden	25-74	211,000	1609	R	90%	50.3 %
Eriksen et al. 1998[12]	Population of Nordic Countries	Finland, Norway, Denmark	-	-	4046	R	~60%[‡]	49.8%
Mullersdorf et al. 2000[17]	Population of Sweden	Sweden	-	-	1422	R	77%	-
Brattberg et al. 1989[8]	A county of Sweden	Sweden	18-84		827	R	82%	53%

[†] If one or more items missing are missing in the table, these items were not presented in this table; [Ψ] R= random sample, NR= non random sample, WP= whole population; [‡] The response rate is estimated by the authors

Table 4.

Method used to obtain information on shoulder and shoulder arm complaints, definition and prevalence rates

First author	Case definition	Outcome measurement	Prevalence overall (95% CI)[Λ]	Prevalence in defined groups (95% CI)[Y]
Shoulder pain				
Point prevalence				
Badley et al. 1992 [7]	Current shoulder pain/swelling or stiffness; 'among anyone in your household'	Questionnaire	6.9% (6.7%-7.1%)	16-24 years 0.7% (-1.7%-4.6%) 25-34 years 2.3% (0.1%-4.6%) 35-44 years 5.5% (3.2-7.8%) 45-54 years 8.2% (5.8%-10.6%) 55-64 years 11.8% (9.4%-14.2%) 65-74 years 12.1% (9.6%-14.6%) 74-85 years 13.2% (9.9-16.5%) 85+ years 15.9% (8.4%-23.4%)
Gomez et al. 1997 [13]	Current shoulder pain for a least one month, and one or more positive signs on physical examination	Interviews, if having pain followed by physical examination	7.8% (6.1%-9.5%)	

First author	Case definition	Outcome measurement	Prevalence overall (95% CI)[Λ]	Prevalence in defined groups (95% CI)[Y]
Allander 1974 [5]	Current shoulder pain and clinically confirmed restricted movements manifested as insufficient passive abduction and inward rotation with concomitant pain. Crepitation and/or painful motion do not qualify	Physical examination	-	31-35 years 8% (6.1-10.0%) 42-46 years 15% 56-60 years 27% (24.6-29.4%) 70-74 years 21%
Adebajo et al. 1992 [4]	Current shoulder pain manifested as: Rotator cuff lesions: diagnosed by criteria of Cyriax Or Frozen shoulder (criteria of Bulgen[β])	Interview, if complaints followed by physical examination	14%[γ] (7.2%-21%)	
Chard et al. 1991 [10]	Current shoulder pain	Interview, when having shoulder pain, followed by physical examination by rheumatologist	26% (22.6%-29.4%)	
One month prevalence				
Urwin et al. 1998 [21]	Shoulder pain > 1 week in previous month	Questionnaire	18.6% (17.6%-19.6%)	16-44 years ¥ 9 % (6.9-11.1%) * 12% (9.6-14.4%) 45-64 years ¥ 19% (16.1-21.9%) * 19% (16.2-21.9%) 65-74 years ¥ 16% (13.3-18.7%)

First author	Case definition	Outcome measurement	Prevalence overall (95% CI)[Λ]	Prevalence in defined groups (95% CI)[Y]
				* 26% (22.8-29.2%) <u>>75 years</u> ¥ 20% (17.1-22.9%) * 24% (20.9-27.1%)
Pope et al. 1997 [19]	Shoulder pain > 24 hours in month before survey	Questionnaire	31% (25.9%-36.1%)	
Makela et al. 1999 [15]	Shoulder pain or tenderness in the previous month	Screening by Questionnaire, interview, functional tests, standard laboratory tests	28% (27.0% - 29.0%)	
<u>One year prevalence</u> Jacobsson et al. 1989 [14]	Subacromial shoulder pain: diagnosed as pure subacromial pain without any pain in the neck (A complaint is only considered when pain or stifness of the joints had been present, continuously or intermittently, for more than 6 weeks during the preceding 12 months).	Interview	4.8 % (2.8%-6.8%)	
Natvig et al. 1994 [18]	Shoulder pain in the previous 12 months	Nordic Questionnaire	46.7% (44.8%-48.6%)	¥ 36.5% (33.9-39.1%) * 56.2% (53.9-59.1%)
Brattberg et al. 1996 [9]	Shoulder problems in the previous 12 months	Interview	34.5% (30.5%-38.5%)	<u>77-79 years</u> ¥ 20.4%

First author	Case definition	Outcome measurement	Prevalence overall (95% CI)[Λ]	Prevalence in defined groups (95% CI)[¥]
				* 34.9% 80-84 years ¥ 30.0% * 45.1% 85+ years ¥ 32.2% * 30.6%
Life time prevalence				
Cunningham et al. 1984[11]	Ever experienced certain symptoms related to arthritis (Joint pain, aching, swelling, morning stiffness, associated disability)	Interview, when having symptoms followed by second interview	6.7% (6.1%-7.3%)	
Reyes et al. 2000[20]	Shoulder pain ever	ILAR/COPCORP	10.1% (6.7%-13.5%)	
Meyers et al. 1982[16]	Shoulder pain ever		66.7% (59.4%-74.0%)	
Incidence of shoulder pain				
Allander 1974[5]	Current shoulder pain and clinically confirmed restricted movements manifested as insufficient passive abduction and inward rotation with concomitant pain. Crepitation and/or painful motion do not qualify	First phase: Questionnaire Second phase: Physical examination		31-35 years 0.9% (6.9-11.1%) 42-46 years 2.5% 56-60 years 1.1% (0.5-1.7%) 70-74 years 1.6%

First author	Case definition	Outcome measurement	Prevalence overall (95% CI)[Λ]	Prevalence in defined groups (95% CI)[Y]
Shoulder-arm pain				
Point prevalence				
Andersson et al.1993[6]	Current a specific shoulder/upper arm pain: persistent or regularly recurrent pain with a duration> 3 months, localisation by manikin	Questionnaire	20% (18.0%-22.0%)	¥ 17.7% (15.1-20.3%) * 22.3% (19.4-25.2%)
One month prevalence				
Eriksen et al. 1998[12]	Pain in arms or shoulder in previous 30 days	Ursin health inventory	33% (31.5%-34.4%)	<30 years ¥ 23.5% * 35.7% 30-39 years ¥ 27.8% * 38.4% >50 years ¥ 34.2% * 47.1%
Pope et al. 1997[19]	Pain or aches in shoulders in month before survey, localised on manikin in three different ways: Shoulder, upper arm Shoulder, upper arm, upper trunk Shoulder, upper arm, upper trunk, neck			34% (28.7%-39.3%) 42% (36.5%-47.5%) 48% (42.5%-53.5%)
Life time prevalence				
Mullersdorf et al. 2000[17]	Shoulder / arm long term pain; limiting	pain scoring lists	8.4% (7.0%-9.8%)	

First author	Case definition	Outcome measurement	Prevalence overall (95% CI)[Λ]	Prevalence in defined groups (95% CI)[Y]
	daily activities ever Long term pain: >3 months or regularly recurrent pain as pain recurring more than once a month and lasting more than 24 hours.			
Brattberg et al. 1989 [8]	Shoulder or arm pain/discomfort	Questionnaire	30.7% (27.6%-33.8%)	<u>≤1 month</u> 4.7% (3.3-6.1%) <u>1-6 months</u> 2.8% (1.7-3.9%) <u>6 months</u> 23.2% (20.3-26.1%)

[Λ] Proportion +/- 1.96*se, with se=√((p(1-p))/n); [Y] If the numbers of the defined groups were not specified, the 95% CI could not be calculated and is not presented in this column; [β]In: Bulgen D, Binder A, Hazleman BL. Frozen shoulder: prospective clinical study with an evaluation of three treatment regimens. Ann Rheum Dis 1984;43:353-60; [γ]The prevalence presented in the study was the combination of both rotator cuff lesions and frozen shoulder. ¥ Prevalence among men;* Prevalence among women.

Incidence of shoulder pain Only Allander[5] provided information on annual incidence of shoulder pain in the general population, reporting 0.9% for those aged 31-35 years, 2.5% for 42-46 years, 1.1% for 56-60 years, and 1.6% for those aged 70-74 years. The study population comprised subjects with no previous episodes of shoulder pain. Subjects became cases if they had clinically confirmed restricted shoulder movements, manifested as insufficient passive abduction and inward rotation with concomitant pain in the shoulder.

4 Discussion

This review describes the results of 17 studies on prevalence and 1 on both incidence and prevalence of shoulder pain in the general population. We found 6.9 - 26% for the point prevalence, 18.6 - 31% for the one-month prevalence, 4.7 - 46.7% for the one-year prevalence and 6.7 - 66.7% for the life-time prevalence of shoulder complaints. Although an increase in prevalence rates could be expected when studies examined period prevalences of greater length, this was not evident from the results of this review. The range within a specific spell of time was that large that there was an overlap in outcome for the point-, period-, and life-time prevalences. This seems mainly due to differences in case definitions. Substantially lower prevalence rates within a specific period of time were found for more detailed case definitions, in terms of duration of complaints and/or presence of limited shoulder motion. High prevalence rates (>30%) were reported when the location for pain was enlarged.[19]

The strong difference in case definition hampered a conclusion on the influence of different factors related to the outcome (sample size, methodological quality, age, and ethnicity). Regarding the sample size it is evident that studies[5-7,11,12,15,17,18,21] with a larger sample size (n >1000) estimate the prevalence more precise. However, this seemed not to influence the level of the prevalence. In two studies with a large

sample size, but different case definition, the point prevalence diverged quite much.[5,7] While on the other hand Pope et al.[19] found, with a small sample size but comparable case definition, similar results as Makela et al.[15]

The influence of methodological quality is difficult to assess due to the difference in case definition. Only for the item random sampling two studies were comparable on the case definition and differed on the item. Makela et al.[15] did not sample randomly, but it seemed not to influence the results compared to Pope et al.[19] Although it is not clear from this review what the influence of valid and reliable instruments is on the outcome, there are several validated methods to assess shoulder complaints, for example the Nordic Questionnaire,[22] the VAS for pain,[23] and the Shoulder disability questionnaire.[24] However only two out of 18 studies used a validated instrument.

Age seemed to increase the prevalence, based on the studies that presented prevalence rates for different age groups within the same case definition.[5,7,21] This is not confirmed by studies, which have used different case definitions. Brattberg et al.[9] found a substantially lower one-year prevalence for people older than 76 years than Natvig and Nassoy[18] found for the general population. Similar differences were found for the point prevalence. There were no comparable studies available for the influence of ethnicity.

Surprisingly only one study was found on incidence of shoulder pain in the general population. Allander[5] focussed on the primary onset of shoulder pain and found annual incidence rates ranging from 0.9% to 2.5% for different age groups. The lack of incidence studies in the general population is possibly due to the high costs involved in longitudinal studies compared with of prevalence studies which can be performed using a cross-sectional design.

This review has certain limitations. We omitted publications in non-scientific journals (such as reports based on governmental databases) because these publications are not available in the computerised literature databases. Thus, we analysed only the scientific literature.

Exclusion of these reports possibly have influenced our outcome, however, it was uncertain if we could obtain all reports world-wide. So including them might have also introduced bias. Furthermore, we could have missed articles that were not published in non-indexed journals; we tried to minimise this bias by citation tracking.

The results of this review support the viewpoint that differing definitions of shoulder pain substantially contribute to the wide range of prevalence rates reported in epidemiological studies. This underlines the importance of carefully defining shoulder pain when undertaking population studies, to allow clear interpretation of the results and comparisons with other data. Furthermore, definitions are required that reconcile the need for precision of symptom specification with the reality of clinical practice.

In conclusion, there are substantial differences in reported prevalences of shoulder pain in the general population. Health professionals and policymakers who estimate the amount of medical care needed and their related costs should be aware of these variations and the underlying reasons for this finding.

5 Acknowledgements

We would like to thank Elske Faber and Ruth Spalburg for their help with translation of a Norwegian and a Spanish paper, respectively.

6 References

1. van der Heijden GJ. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
2. Bjelle A. Epidemiology of shoulder problems. Baillieres Clin Rheumatol 1989;3:437-51.
3. Green S, Buchbinder R, Glazier R, Forbes A. Interventions for shoulder pain. Cochrane Database Syst Rev 2000:CD001156.
4. Adebajo AO, Hazleman BL. Soft tissue shoulder lesions in the African. Br J Rheumatol 1992;31:275-6.
5. Allander E. Prevalence, incidence, and remission rates of some common rheumatic diseases or syndromes. Scand J Rheumatol 1974;3:145-53.

6. Andersson HI, Ejlertsson G, Leden I, Rosenberg C. Chronic pain in a geographically defined general population: studies of differences in age, gender, social class, and pain localization. Clin J Pain 1993;9:174-82.
7. Badley EM, Tennant A. Changing profile of joint disorders with age: findings from a postal survey of the population of Calderdale, West Yorkshire, United Kingdom. Ann Rheum Dis 1992;51:366-71.
8. Brattberg G, Thorslund M, Wikman A. The prevalence of pain in a general population. The results of a postal survey in a county of Sweden. Pain 1989;37:215-22.
9. Brattberg G, Parker MG, Thorslund M. The prevalence of pain among the oldest old in Sweden. Pain 1996;67:29-34.
10. Chard MD, Hazleman R, Hazleman BL, King RH, Reiss BB. Shoulder disorders in the elderly: a community survey. Arthritis Rheum 1991;34:766-9.
11. Cunningham LS, Kelsey JL. Epidemiology of musculoskeletal impairments and associated disability. Am J Public Health 1984;74:574-9.
12. Eriksen HR, Svendsrod R, Ursin G, Ursin H. Prevalence of subjective health complaints in the Nordic European countries in 1993. European Journal of Public Health 1998;8:294-8.
13. Gomez MR, Diz PG, Barros GL, Gomez CS, Lozano Eire MJ, Robisco LP et al. Estudio epidemiologico de la patologia dolorosa del hombro en nuestro medio. Rev Esp Reumatol 1997;24:247-50.
14. Jacobsson L, Lindgarde F, Manthorpe R. The commonest rheumatic complaints of over six weeks' duration in a twelve-month period in a defined Swedish population. Prevalences and relationships. Scand J Rheumatol 1989;18:353-60.
15. Makela M, Heliovaara M, Sainio P, Knekt P, Impivaara O, Aromaa A. Shoulder joint impairment among Finns aged 30 years or over: prevalence, risk factors and co-morbidity. Rheumatology (Oxford) 1999;38:656-62.
16. Meyers OL, Jessop S, Klemp P. The epidemiology of rheumatic disease in a rural and an urban population over the age of 65 years. S Afr Med J 1982;62:403-5.
17. Mullersdorf M, Soderback I. The actual state of the effects, treatment and incidence of disabling pain in a gender perspective-- a Swedish study. Disabil Rehabil 2000; 22:840-54.
18. Natvig B, Nassioy I. Musculoskeletal complaints in a population. Occurrence . Tidsskrift for Den Norske Laegeforening 1994;114:323-7.
19. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: the influence of case definition. Ann Rheum Dis 1997;56:308-12.
20. Reyes Llerena GA, Guibert Toledano M, Hernandez Martinez AA, Gonzalez Otero ZA, Alcocer Varela J, Cardiel MH et al. Prevalence of musculoskeletal complaints and disability in Cuba. A community-based study using the COPCORD core questionnaire. Clin Exp Rheumatol 2000;18:739-42.
21. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M et al. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of

symptoms at different anatomical sites, and the relation to social deprivation. Ann Rheum Dis 1998;57:649-55.

22. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sorensen F, Andersson G, et al. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 1987;18:233-7.
23. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983;17:45-56.
24. Van der Heijden GJ, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. J Clin Epidemiol 2000;53:29-38.

Chapter 3

High incidence and recurrence of shoulder and neck pain in a working population during a 2 year follow-up

Accepted for publication by the Journal of Clinical Epidemiology

Abstract

Background and Objective Little is known about the long term course of shoulder and neck complaints. The objective of this study is to describe the course of these complaints in a working population over time. **Methods** A longitudinal study with 2 follow-up measurements was performed among workers of home for the elderly and care homes. Annual questionnaires were administered on neck and shoulder complaints and descriptive statistics were used to analyse these data. **Results** 12-month incidence rates for neck and shoulder complaints of 16%-18% were observed, 12-month prevalence rates roughly twice as high, and 12-month recurrence rates approximately twice the prevalence rates. Each year medical care was sought by 21%-38% of the subjects with neck or shoulder pain and 13%-21% were absent to work. Although at population level the occurrence of neck and shoulder complaints remained constant, the course of complaints within individuals demonstrated a strong episodic nature of neck and shoulder pain. **Conclusion** Results from this study suggest that neck and shoulder complaints for most subjects runs a recurrent course characterised a strong variation in occurrence, rather than an acute, self-limiting course. These findings suggest that clinical trials should have a sufficiently long follow-up period to demonstrate sustainability of the therapeutic results.

1 Introduction

Shoulder and neck pain are common problems in the general population with one-year prevalences varying between 7-47% for shoulder complaints[1,2] and 16-61% for neck complaints.[3] Shoulder and neck complaints may result in sick leave, loss of productivity, and inability to carry out household and leisure time activities.[4,5] In the period 1987-1995 Washington State paid over 66 million dollars a year on work disability claims due to shoulder problems.[6] Swedish insurance data on income compensation during sick leave showed that in 1994 the costs for neck-shoulder complaints equalled the costs for back pain.[7]

Despite the societal impact of these complaints, little is known about the long-term course of shoulder and neck complaints due to the lack of longitudinal studies.[8,9] Prospective studies among patients with shoulder complaints in primary care have shown that 41%-51% of these patients experienced recurrent episodes during an 18-month follow-up.[4,10] In one cohort study among patients with shoulder pain more than 50% of all subjects still reported shoulder pain with disabling symptoms about 3 year later.[11] Comparable studies on the occurrence of neck pain over time are lacking. The prospective studies suggest that shoulder complaints are quite persistent[11] or, alternatively, vary considerably over time with fluctuating severity of these complaints.[10]

To understand the natural course of neck and shoulder complaints and their impact on daily life, we studied the dynamic patterns of incidence, recurrence, chronicity, and recovery of shoulder and neck complaints in a longitudinal cohort study with 2 years follow-up. The aims of this study were (i) to describe the temporal changes in incidence, recurrence, chronicity, and recovery of neck and shoulder complaints over a three year period, (ii) to identify the characteristics of pain predicting the recurrence, and (iii) to evaluate the effect of the characteristics of neck and shoulder complaints on care seeking behaviour.

2 Methods

Population The current study is part of a longitudinal study with two-year follow-up on the development of musculoskeletal complaints in an occupational population. The source population consisted of all subjects working in 4 nursing homes and 3 homes for the elderly in the Netherlands. The health care sector is well known for its high proportion of musculoskeletal complaints among its personnel.[12] Subjects were invited to participate in the study if they had worked for more than 12 months for over 10 hours per week in their current job. Between March 1998 and March 1999 self-administrated questionnaires were distributed. Follow-up measurements among respondents were carried out 1 and 2 years later, using a similar questionnaire.
At baseline 1208 workers were eligible to enroll in the study, of which 769 (64%) responded. The responders worked in a variety of professions, such as nurse (n=129), care giver (n=264), kitchen worker (n=58), housekeeper and cleaner (n=49), maintenance worker (n=14), (physical) therapy (n=38), office work (n=146), and various other jobs (n=62). At 1-year follow-up 529 (68%) of the subjects filled out the questionnaire again, and at 2-year follow-up 346 (65%) subjects returned the questionnaire.

Data collection A questionnaire was used to collect personal data, details on the respondent's job, employment history, general health status, leisure time, and the presence of shoulder complaints. Information on individual factors like age, height, weight, education level and employment in current job was derived from a standardised questionnaire.[13] In addition; a measure of perceived general health was constructed based upon 11 dichotomised questions on general health, excluding musculoskeletal symptoms. This scale has a good internal reliability (Cronbach's α = 0.86) and test-retest reliability (Pearson's r = 0.76).[14] A sum score was calculated over all 11 items and a classification of poor/fair general health was assigned to respondents who scored in the lowest half of the study population.

The questions on neck and shoulder complaints were derived from the standardised Nordic questionnaire for musculoskeletal symptoms. This questionnaire has been shown to be a valid instrument to collect information on the nature, duration (days) and frequency (occurrences) of musculoskeletal symptoms.[15] Subjects were presented a drawing with a pre-shaded area (see supplement) indicating the shoulder area and asked whether they had experienced pain or discomfort which lasted for at least a few hours during the past 12 months. Subsequent questions related to the duration and frequency of these complaints and to periods of sickness absence due to these complaints.[16] Similar questions were asked for complaints in the neck region, again using a pre-shaded area to define the neck. The pre-shaded shoulder and neck region did not overlap.

Care seeking was registered by additional questions about medical consultation in the previous 12 months.[17] A distinction was made between consulting a general practitioner, a physiotherapist, or a medical specialist. All medical specialists, including orthopaedists, surgeons, and neurologists, were grouped under speciality medical care.

Definition of cases with complaints Two definitions of shoulder complaints were used: (a) shoulder complaints in the past 12 months referred to at least one episode of pain or discomfort in the past 12 months for at least a few hours, (b) chronic shoulder complaints in the past 12 months referred to pain or discomfort which was present almost every day in the preceding 12 months with a minimal presence for at least 3 months.[15] Cases with chronic shoulder complaints were by definition a subgroup of all cases with shoulder complaints. Neck complaints were defined in the same way and independently of the presence of shoulder complaints. Hence, shoulder and neck complaints could occur simultaneously in the same case.

We used information on the 12-month prevalence, incidence and recurrence to obtain information on the clinical course of neck and

shoulder complaints. A prevalent case was defined as a subject having had at least one episode of complaints during the previous 12 months. These cases were determined during the baseline survey, follow-up 1, and follow-up 2. An incident case was defined as a subject, which experienced a new episode of complaints during 12 months after at least 12 months free of these complaints. Recurrent cases were subjects, which experienced episodes of complaints in subsequent years. Thus, incident and recurrent cases could be identified during follow-up 1 and follow-up 2. Given these definitions, a subject was regarded as recovered when a year with complaints was followed by a year free of complaints.

Statistical analysis Descriptive techniques were used to present the temporal changes of shoulder and neck complaints over time. In order to avoid different denominators during the follow-up due to loss-to-follow-up, the dynamic changes in presence and severity of shoulder and neck complaints and subsequent care seeking were described among subjects who completed all 3 consecutive questionnaires (n=346). A non-response analysis was conducted to analyse whether the annual dropout biased the results.

Analytic techniques were used to analyse the associations between complaint characteristics and recurrence of complaints in the following year. These associations were expressed by relative risks (RR) with 95% confidence intervals. Data of care seeking behaviour was analysed cross-sectional. Both episode and related visit to a caretaker occurred in the same year, without having information which episode of complaints in a given year prompted health care seeking behaviour. These associations were expressed by odds ratio (OR) with 95% confidence intervals. All calculations were carried out in the statistical package SAS (Version 6.12).[18]

3 Results

Population Characteristics of the study population at baseline and both follow-up measurements are presented in table 1. The non-responders in follow-up 1 (n=240) and follow-up 2 (n=183) were not significantly different from the responders at baseline for age, height, weight, duration of employment, and occurrence of shoulder or neck complaints.

Table 1

Baseline characteristics of personnel of nursing homes and homes for the elderly, participating in a longitudinal study with 2 year follow-up

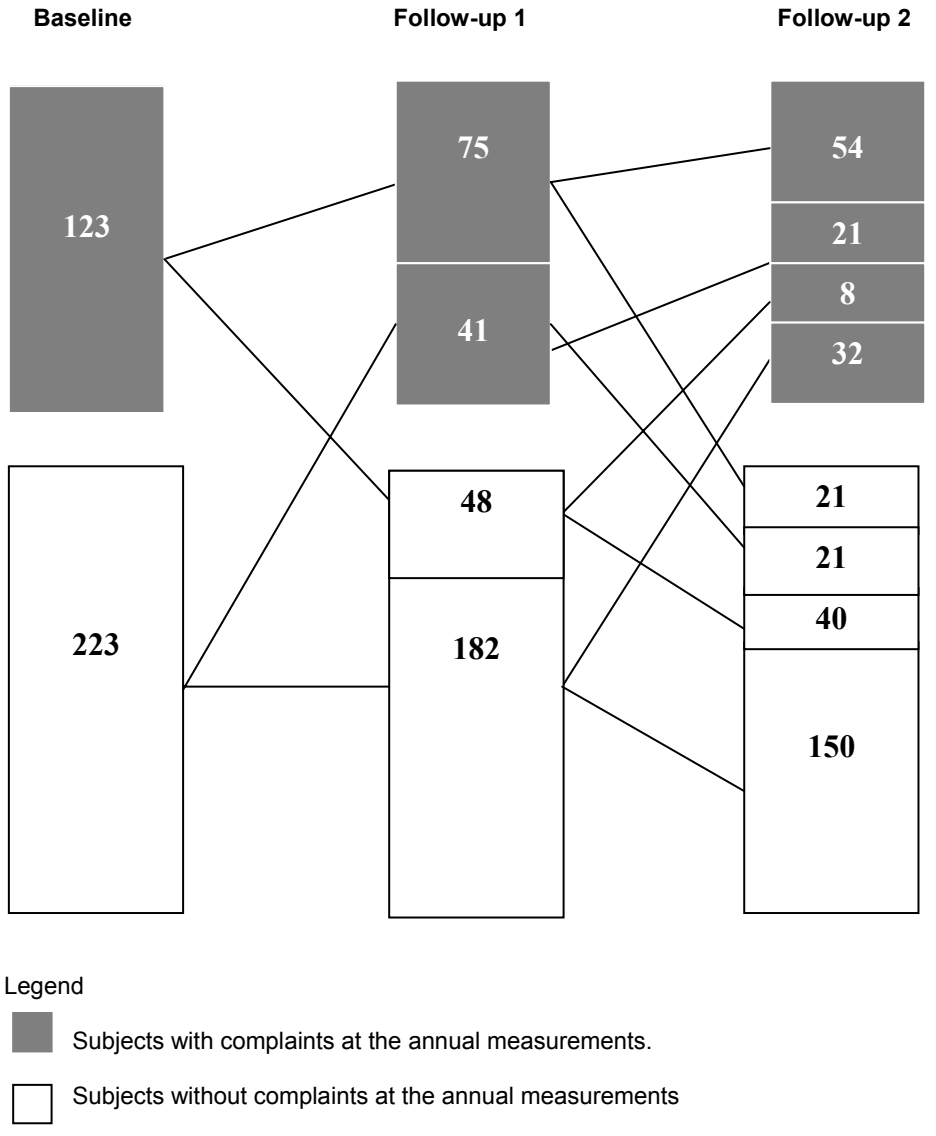
	Baseline(n=769)		Follow-up 1 (n=529)		Follow-up 2 (n=346)	
Sex (% women)	83.9%		85.1%		84.1%	
	Mean	*Stdev*	*Mean*	*Stdev*	*Mean*	*Stdev*
Age (y)	40.0	10.0	40.8	9.7	41.6	8.9
Height (cm)	168.8	9.5	169.1	9.0	169.2	9.0
Weight (cm)	70.9	13.6	71.3	13.7	71.5	14.0
Employment in current job (y)	8.0	6.7	9.3	6.9	9.5	6.6

Course of neck and shoulder complaints over time The individual course of the subjects with neck or shoulder pain, which are shown in figure 1 and 2, showed a dynamic pattern, while the prevalence remained stable over the 3 year period. At baseline, 123 out of 346 (35.6%) had experienced neck complaints in the past 12 months. During the first year of follow-up 75 subjects (61.0%) again reported episodes of neck complaints. Among those initially free of symptoms 41 subjects (18.4%) experienced neck complaints during the first year of follow-up, whereas 182 subjects (81.6%) remained free of complaints. A similar pattern was observed during the second year of the follow-up. In figure 2 it is depicted that the course of shoulder complaints over time showed a similar pattern as neck complaints. During the 3 consecutive annual measurements 54 workers (15.6%) consistently reported the presence of neck complaints in the past year whereas 150 workers

(43.4%) consistently reported the absence of these complaints. These proportions for shoulder complaints were 16.8% and 46.5%, respectively.

Figure 1

The course of neck complaints during a three year period among workers with 3 consecutive measurements in a longitudinal study with 2 year follow-up (n=346)

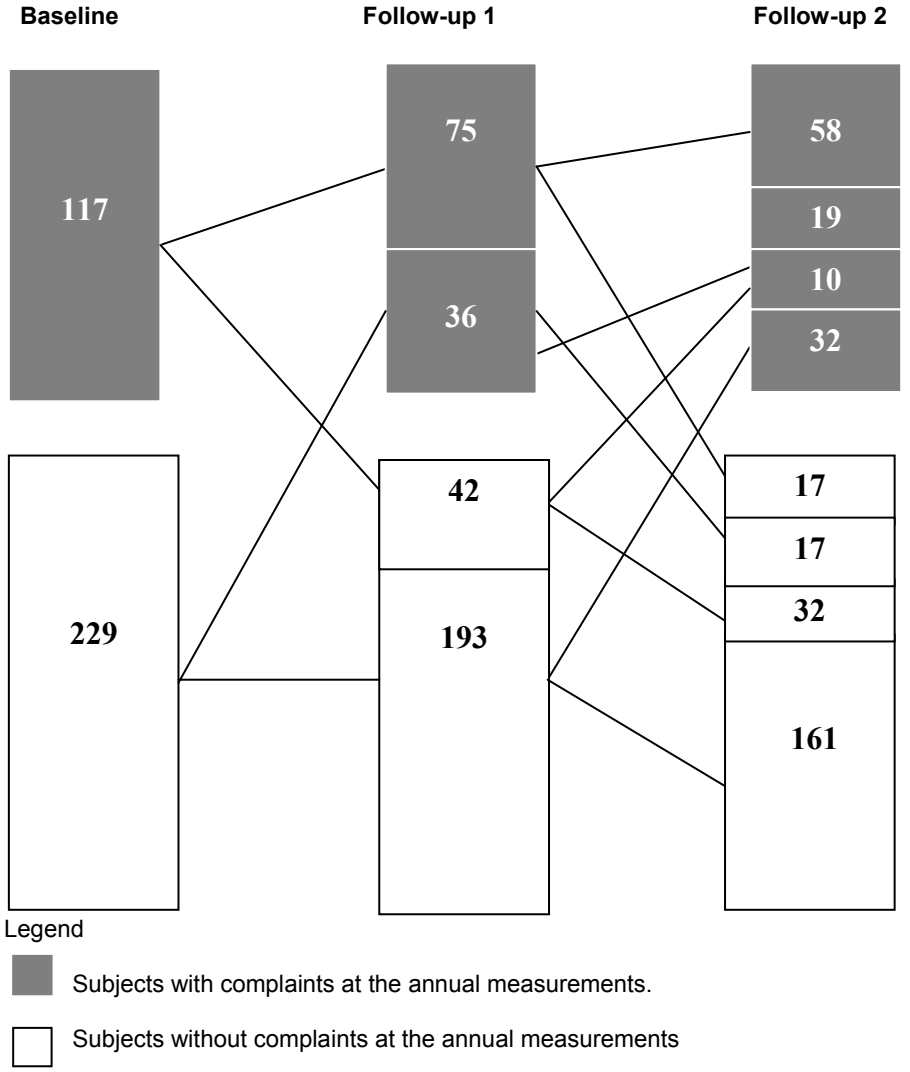


In addition, both figures include information on the composition of the 12-month prevalence. In the latter block of figure 2 is shown that there were 119 prevalent cases at follow-up 2, divided over four groups. These groups included 58 subjects who had recurrent complaints in

both follow-ups and 19 subjects with recurrent neck pain who were at baseline free of neck pain**.** Another 42 subjects had incident neck pain of which 32 subjects were free of neck pain at baseline and first follow-up and 10 subjects had prevalent complaints at baseline, but were free of complaints at first follow-up.

Figure 2

The course of shoulder complaints during a three year period among workers with 3 consecutive measurements in a longitudinal study with 2 year follow-up (n=346)



Prevalence, incidence and recurrence The annual values of prevalence, incidence, and recurrence of neck and shoulder complaints are presented in Table 2. The 12-month prevalence for neck complaints varied between 33% and 36%. During the first and second year of follow-up the incidence and recurrence rates were 17% and 18%, and 61% and 65%, respectively. For shoulder complaints similar results were found with 12-month prevalences between 32% and 34%, incidence rates of 16% and 18%, and recurrence rates of 64% and 69%. The 12-month prevalence and incidence of complaints with duration longer than 3 months (chronic complaints) consistently showed a proportion of 20-30% of neck or shoulder complaints. Recurrence of these chronic complaints varied between 28% and 57% for neck complaints and between 31% and 66% for shoulder complaints. When considering any complaint episode as a recurrent event, whether or not of chronic nature, the recurrence rates for chronic neck complaints were 73% and 80% and for shoulder complaints 78% and 83%.

Table 2

Prevalence, incidence, and recurrence of shoulder and neck complaints among personnel of nursing homes and homes for the elderly with 3 consecutive measurements in a longitudinal study with 2 year follow-up (n=346)

	Baseline (95% CI)	Follow-up 1 (95% CI)	Follow-up 2 (95% CI)
Neck complaints in past 12 months			
Prevalence	35.6% (30.6%-40.7%)	33.5% (28.5%-38.5%)	33.2% (28.2%-38.2%)
Incidence	-	18.4% (13.3%-23.5%)	17.4% (12.5%-22.3%)
Recurrence	-	61.0% (54.6%-67.4%)	64.7% (58.5%-70.9%)
Neck complaints present at least 3 months in past 12 months			
Prevalence	8.7% (5.7%-11.7%)	8.7% (5.7%-11.7%)	10.1% (6.9%-13.3%)
Incidence	-	6.6% (3.3%-9.9%)	5.7% (2.7%-8.7%)
Recurrence	-	30.0%	56.7%

	Baseline (95% CI)	Follow-up 1 (95% CI)	Follow-up 2 (95% CI)
		(22.0%-36.0%)	(50.3%-63.1%)
Shoulder complaints in past 12 months			
Prevalence	33.8%	32.1%	34.4%
	(28.8%-38.8%)	(27.1%-37.0%)	(29.4%-39.4%)
Incidence	-	15.7%	17.9%
		(11.0-20.4%)	(13.0%-22.8%)
Recurrence	-	64.1%	69.4%
		(57.9%-70.3%)	(63.5%-75.3%)
Shoulder complaints present at least 3 months in past 12 months			
Prevalence	9.3%	8.4%	10.4%
	(6.2%-12.4%)	(5.5%-11.3%)	(7.2%-13.6%)
Incidence	-	6.1%	5.4%
		(3.0%-9.2%)	(2.5%-8.3%)
Recurrence	-	31.3%	65.5%
		(25.3%-37.1%)	(59.4%-71.6%)

Medical care seeking A substantial proportion of the subjects with neck or shoulder complaints sought medical care (see Table 3). Among workers with neck pain, 21%-30% yearly visited their general practitioner (GP) and almost a similar proportion sought care through a physiotherapist. Consultation of a medical specialist occurred less frequently (3%-7%). A comparable pattern was observed for medical care seeking among workers with shoulder complaints, although the percentages of workers seeking care were slightly higher. Visiting a GP because of neck complaints was associated at baseline with the presence of the complaint longer than 3 months (OR 5.84; 95% CI 2.51-13.61) and sickness absence due to the complaint (OR 4.18; 95% 1.52-11.49). Similar associations were found for shoulder complaints and for visiting one of the other caretakers, such as medical specialists. Age, sex, and education did not influence care seeking although a tendency was observed for subjects with lower and intermediate education to seek care at their general practitioner more often than higher educated subjects with complaints.

Table 3

Care seeking among subjects with neck or shoulder complaints in the past 12 months among personnel of nursing homes and homes for the elderly with 3 consecutive measurements in a longitudinal study with 2 year follow-up (n=346)

	Baseline (95% CI)	Follow-up 1 (95% CI)	Follow-up 2 (95% CI)
Neck complaints	n=123	n=116	n=115
General practitioner	30.1% (22.0%-38.1%)	25.0% (17.1%-32.9%)	20.9% (13.5%-28.3%)
Medical specialist	7.3% (2.7%-11.9%)	5.2% (1.2%-9.2%)	2.6% (0.0%-5.5%)
Physiotherapist	26.8% (19.0%-34.6%)	20.7% (13.3%-28.1%)	21.7% (14.2-29.2%)
Shoulder complaints	n=117	n=111	n=119
General practitioner	37.6% (28.8%-46.3%)	36.0% (27.1%-45.0%)	24.4% (16.7%-32.1%)
Medical specialist	11.1% (5.4%-16.8%)	9.0% (3.7%-14.3%)	5.0% (1.1%-8.9%)
Physiotherapist	29.1% (20.9%-37.3%)	31.5% (22.9%-40.1%)	22.7% (15.2-30.2%)

Sickness absence due to shoulder pain was present in 13.7% (n=16) of all subjects with shoulder complaints at baseline. During the two follow-up measurements these proportions were 24.3% (n=27) and 17.7% (n=21). The findings for sickness absence due to neck complaints were comparable with proportions of 13.8% (n=17), 19.0% (n=22), and 18.3% (n=21), respectively.

Risk factors for recurrence of complaints are illustrated in Table 4. Subjects with complaints lasting longer than 3 months in a given year had a significantly increased risk on recurrence of complaints in the following year. The relative risk varied between 2.31 and 3.56 for neck complaints and between 2.50 and 4.06 for shoulder complaints. A subject who experienced more than 5 episodes of neck or shoulder complaints during a year also had a higher probability on recurrence of these complaints in the next year.

Although of lesser importance, a poor/fair general health was a significant predictor for recurrence of neck or shoulder complaints in the next year. Other factors, such as age, sex, education, and duration of employment did not play a significant role. In multivariate analyses both chronic and frequently occurring complaints remained significant predictors, whereas the influence of poor/fair general health did not reached statistically significant levels ($p<0.05$).

Table 4

Univariate analysis of the associations between symptoms and general health with recurrence of neck and shoulder complaints in the next year among personnel of nursing homes and homes for the elderly

Factor	Follow-up 1 (n=529) Relative Risk (95% CI)	Follow-up 2 (n=346) Relative Risk (95% CI)
Neck complaints		
Complaints present longer than 3 months in past year	2.31* (1.69 – 3.15)	3.56* (2.30 – 5.52)
More than 5 episodes in past year	2.67* (2.10 – 3.40)	2.93* (2.22 – 3.88)
Poor/fair general health	1.26* (1.12 – 1.43)	1.24* (1.06 – 1.45)
Shoulder complaints		
Complaints present longer than 3 months in past year	3.33* (2.40 – 4.61)	4.06* (2.51 – 6.57)
More than 5 episodes in past year	2.50* (1.96 – 3.19)	3.47* (2.49 – 4.85)
Poor/fair general health	1.34* (1.19 – 1.52)	1.27* (1.09 – 1.48)

* Mantel-Haenszel test, $p < 0.05$

4 Discussion

In this longitudinal study in a working population we observed 12-month incidence rates for neck and shoulder complaints of 16%-18%, 12-month prevalence rates roughly twice as high, and 12-month recurrence rates approximately twice the prevalence rates. The concurrence of both neck and shoulder complaints was reflected by the fact that 50%-60% of all subjects with neck complaints also reported to have experienced shoulder complaints in the past 12 months. The course of neck and shoulder complaints over time demonstrated a highly dynamic pattern in the occurrence of complaints. Medical care seeking among subjects with neck or shoulder complaints was high, with approximately 21%-38% visiting a general practitioner and a physiotherapist each year. Complaint characteristics largely determined care-seeking behaviour.

The initial response rate among workers was 64%, which may have given rise to selection bias. We observed response rates between 48% and 82% in the participating 7 organisations, but the response rates were not associated with the observed prevalences of neck and shoulder complaints. This result suggests that the reported occurrence of complaints was not substantially influenced by selective participation.

A remark has to be made on the finding that the recurrence rates of chronic complaints of neck or shoulder in the second follow-up were twice as high as those are during the first year of follow-up. An explanation for this difference is difficult, since some selection must have occurred with under representation of chronic cases during the first follow-up. However, when estimating the probability of any recurrence of complaints for chronic complaints in the baseline survey, no differences were observed between both years of follow-up. It is difficult to appreciate how this may have affected the results, since determinants of care seeking and predictors for recurrent complaints were very similar across the total follow-up period. Another longitudinal study[19] on chronic pain demonstrated that 79% of the subjects with

chronic pain at baseline still have chronic pain four years later. Although this study comprehended not only musculoskeletal pain, it showed clearly the high recurrence of chronic pain.

There are several aspects of the case definition, which have or might have had influence on the outcome of our study. Firstly, it has been shown that different case definitions of shoulder pain lead to varying prevalence estimates of the disorder.[20] We have used drawings with a restricted area for the shoulder and neck, which most likely will decrease the estimated prevalence. On the other hand, our definition of complaint was based on pain or discomfort lasting for at least a few hours, which will certainly have increased the estimated prevalence when compared to more restrict definitions of pain and the presence of disability.[20] In a substantial number of cases with neck or shoulder complaints, movements of the neck and shoulder can provoke the symptoms

Secondly, both incidence and recurrence of neck and shoulder complaints were determined by the recall period of 12 months. A disadvantage of this long recall period is that subjects may have experienced several complaint episodes within this year and, thus, an incident case may already have had recurrent episodes. A shorter recall period will undoubtedly increase the variability in presence or absence of neck and shoulder complaints that it may even become impossible to separate between two episodes.

A third aspect is the case definition of neck and shoulder complaints. We used the Nordic Questionnaire, as shown in supplement I, which differentiates between neck and shoulder complaints. This approach does not distinguish between shoulder pain and neck pain referring in the shoulder region. The large overlap between neck and shoulder pain in this study suggests that to some extent both complaints may stem from the same underlying clinical condition in one region with referred pain to the other region. Therefore, one might argue that the separation between both neck and shoulder complaints is partly arbitrarily. A large overlap between neck and shoulder complaints was also observed by

Viikari-Juntura and colleagues among workers seeking medical advice from an occupational health service.[21] However, a further classification into local neck pain, shoulder pain, and neck pain with coexisting numbness or pain in the forearm or hand did not have any predictive value for sick leave due to neck and shoulder disorders. We preferred to describe both complaints separately in order to be able to distinguish between incident and recurrent complaints in the same body region. However, it has to be acknowledged that some cases with neck or shoulder complaints are misclassified and in fact the result of referred pain.

The 12-month prevalence of neck or shoulder complaints in our study cohort was in the range of observations in cross-sectional surveys. In the general population of Norway and the United Kingdom 12-month prevalence rates for neck pain of respectively 35%, 33%, and 43% have been reported.[22-24] Using the same Nordic questionnaire, another survey in Norway estimated a 12-months prevalence of shoulder pain in the general population of 47%, which is higher than in our working population.[2] The incidence of neck pain in our study was 18%, measured over a recall period of 12 months. This is very similar to the 1-year incidence of neck complaints of 18% in the general population[25] and 20% in a working population.[26] Few reports on recurrence rates have been published, but our recurrence rates of over 60% within 12 months are higher than the recurrence rates of 41%-51% among patients with shoulder complaints in primary care during a 18-month follow-up.[4,10,11]

Our study showed that a substantial proportion of the subjects with neck and shoulder complaints had recurrent episodes of these complaints in the next year of follow-up. In another study among health care workers it was also observed that about 50% of the subjects with musculoskeletal complaints varied between being a case or not during a three year period.[27] Previous studies showed that a history of persistent shoulder complaints is a good predictor for slow recovery over time and/or recurrence of complaints, similar to other regional pain

syndromes.[10] Our study supports these findings and illustrates that for most subjects neck and shoulder pain are not characterised by a single episode in time. More often it is likely to be a flare up of an episodic condition, as has been described for back pain by several authors.[28-31] The finding of the high recurrence rate (>60%) of neck and shoulder complaints may have consequences for further studies. Firstly, cohort studies on risk factors for the occurrence of neck and shoulder complaints should not only pay attention to new episodes, but also analyse risk factors for recurrent episodes (flare-ups). In analysing recurrent episodes the question is not whether an episode occurs but when an episode occurs, and which factors are responsible for triggering this event. A time-related question like this one is not easy to answer with a normal cohort study design, but would require a case-crossover analysis within a cohort study, as described by Mittleman et al.[32] Prospectively data should be gathered about stable, intermittent risk factors and complaints. Risk patterns for new episodes are analysed the usual way, while for recurrent episodes information is collected on intermittent risk factors immediately preceding the episode. In using this approach, each individual forms his or her own stratum. An alternative could be to study the frequency of complaints in a certain time window rather than defining incidence or recurrence in consecutive time windows.

Secondly, it should be considered to increase the follow-up period in controlled trials on intervention for neck and shoulder complaints. Normally, these studies have short follow-up periods of 6 to 12 months. Regarding the high recurrence rate, a follow-up of 18 to 24 months is recommendable, which would also demonstrate the sustainability of the therapeutic results.

In conclusion: neck and shoulder complaints were characterised by temporal changes manifested in high rates of incidence, recurrence, and recovery. Although at population level the occurrence of neck and shoulder complaints remained constant, the course of complaints within individuals demonstrated the strong episodic nature of complaints of

neck and shoulder pain. Recurrence of complaints was predicted by long lasting or frequent episodes of complaints. Care seeking was present in one third of the subjects with complaints and associated with long lasting or frequent complaints and sickness absence.

5 Acknowledgements

We would like to thank Ingrid J.M. Hendriksen, PhD and Arianne P. Verhagen, PhD for their critical suggestions on this article.

6 References

1. Jacobsson L, Lindgarde F, Manthorpe R. The commonest rheumatic complaints of over six weeks' duration in a twelve-month period in a defined Swedish population. Prevalences and relationships. Scand J Rheumatol 1989;18:353-60.
2. Natvig B, Nassioy I. Musculoskeletal complaints in a population. Occurrence. Tidsskr Nor Laegeforen 1994;114:323-7.
3. Nachemson AL, Waddell H, Norlund AI. Epidemiology of neck and low back pain. In: Nachemson AL, Jonsson E. Neck and Back Pain. The Scientific Evidence of Causes, Diagnosis, and Treatment. Philadelphia USA: Lippincott Williams & Wilkins, 2000.
4. Croft P, Pope D, Silman. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
5. Trinkoff AM, Lipscomb JA, Geiger-Brown J, Brady B. Musculoskeletal problems of the neck, shoulder, and back and functional consequences in nurses. Am J Ind Med 2002;41:170-8.
6. Silverstein B, Welp E, Nelson N, Kalat J. Claims incidence of work-related disorders of the upper extremities: Washington State, 1987 through 1995. Am J Pub Health 1998;88:1827-33.
7. Nygren A, Berglund A, von Koch M. Neck-and-shoulder pain, an increasing problem. Strategies for using insurance material to follow trends. Scand J Rehab Med 1995;32(suppl):107-12.
8. Van der Windt DA, Thomas E, Pope DP, de Winter AF, Macfarlane GJ, Bouter LM, Silman AJ. Occupational risk factors for shoulder pain: a systematic review. Occup Environ Med 2000;57:433-42.
9. Borghouts JA, Koes BW, Bouter LM. The clinical course and prognostic factors of non-specific neck pain: a systematic review. Pain 1998;77:1-13.
10. Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-de Jong B. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology

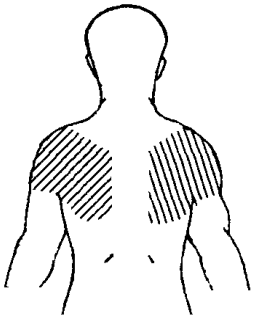
(Oxford) 1999;38:160-3.

11. Macfarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. J Rheumatol 1998;25:1612-5.
12. Lagerström M, Wenemark M, Hagberg M, Hjelm EW. Occupational and individual factors related to musculoskeletal symptoms in five body regions among Swedish nursing personnel. Int Arch Occup Environ Health 1995;68:27-35.
13. Central Bureau of Statistics. Netherlands health interview survey 1981-1997. SDU/publishers, The Hague, 1992.
14. Van Sonsbeek JLA. The Voeg: A list of subjective health complaints. SDU/ publishers (*Statistical reports M37*), The Hague, 1990.
15. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sorensen F, Andersson G, Jorgensen K. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 1987;18:233-7.
16. Burdorf A, Post W, Bruggeling T. Reliability of a questionnaire on sickness absence with specific attention to absence due to back pain and respiratory complaints. Occup Environ Med 1996;53:58-62.
17. Molano SM, Burdorf A, Elders LAM. Factors associated with medical care-seeking due to low back pain in scaffolders. Am J Ind Med 2001;40:275-81.
18. SAS Institute Inc. SAS/STAT® Software: Changes and enhancements through release 6.12. Cary, NC: SAS institute Inc,1997
19. Elliott AM, Smith BH, Hannaford PC, Smith WC, Chambers WA. The course of chronic pain in the community : results of a 4-year follow-up study. Pain 2002;99:299-307.
20. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: the influence of case definition. Ann Rheum Dis 1997;56:308-12.
21. Viikari-Juntura E, Takala E, Riihimaki H, Martikainen R, Jappinen P. Predictive validity of symptoms and signs in the neck and shoulders. J Clin Epidemiol 2000;53:800-8.
22. Bovim G, Schrader H, Sand T. Neck pain in the general population. Spine 1994;19:1307-9.
23. Palmer KT, Walker-Bone K, Griffin MJ, Syddall H, Pannett B, Coggon D, Cooper C. Prevalence and occupational associations of neck pain in the British population. Scand J Work Environ Health 2001;27:49-56.
24. Guez M, Hildingsson C, Nilsson M, Toolanen G. The prevalence of neck pain: a population-based study from Northern Sweden. Acta Orthop Scand 2002;73:455-9.
25. Croft PR, Lewis M, Papageorgiou AC, Thomas E, Jayson MI, Macfarlane GJ, Silman AJ. Risk factors for neck pain : a longitudinal study in the general population. Pain 2001;93:317-25.
26. Leclerc A, Niedhammer I, Landre MF, Ozguler A, Etore P, Pietri-Taleb F. One-year predictive factors for various aspects of neck disorders. Spine 1999;24:1455-62.
27. Josephson M, Lagerström M, Hagberg M, Wigaeus-Hjelm E. Musculoskeletal symptoms and job strain among nursing personnel: a study over a three year period. Occup Environ

Med 1997;54:681-5.
28. Von Korff M. Studying the natural history of back pain. Spine 1994;19:2041S-6S.
29. Von Korff M, Saunders K. The course of back pain in primary care. Spine 1996;21:2833-37.
30. Smedley J, Inskip H, Cooper C, Coggon D. Natural history of low back pain. A longitudinal study in nurses. Spine 1998;23:2422-6.
31. De Vet HC, Heymans MW, Dunn KM, Pope DP, van der Beek AJ, Macfarlane GJ, et al. Episodes of low back pain: a proposal for uniform definitions to be used in research. Spine 2002;27:2409-16.
32. Mittleman MA, Maldonado G, Gerberich SG, Smith GS, Sorock GS. Alternative approaches to analytical designs in occupational injury epidemiology. Am J Ind Med. 1997;32:129-41

Supplement

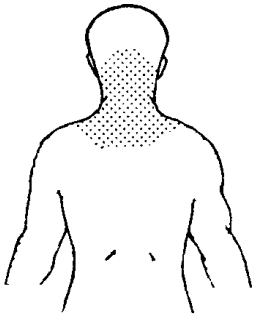
Preshaded manikins used in the Nordiq Questionnaire, with accompanying text as used in the questionnaire.



Shoulder

In this picture you can see the part of the body referred to in this questionnaire as shoulder. By shoulder problems are meant: pain, discomfort, stiffness or numbness in the shaded area. There are separate questions on neck complaints.

Neck



In this picture you can see the part of the body referred to in this questionnaire as neck. By neck problems are meant: pain, discomfort, stiffness or numbness in the shaded area, independent of adjacent areas. There are separate questions on shoulder complaints.

Chapter 4

Work-related risk factors for incidence and recurrence of shoulder and neck complaints in nursing home and elder care workers

Published as: Work-related risk factors for incidence and recurrence of shoulder and neck complaints in nursing home and elder care workers. Jolanda J. Luime, Judith I. Kuiper, Bart W. Koes, Jan A.N. Verhaar, Harald S. Miedema, Alex Burdorf Scand J Work Environ Health 2004;30:279-286

Abstract

Objective To assess the differences and similarities in work related physical, psychosocial and personal risk factors for the incidence and for the recurrence of neck and shoulder complaints. **Methods** A prospective cohort study of 769 workers of nursing homes and homes for the elderly was conducted. At baseline data were collected by questionnaire on personal characteristics, physical work load, psycho-social work load, and the presence of shoulder and neck complaints. After 1 and 2 year follow-up data were collected on shoulder and neck complaints. Generalized estimation equations were used for analyzing risk factors among subjects with at least one follow-up measurement available (n=556, 72%). **Results** In the multivariate model, adjusted for age and gender, obesity (OR 2.12; 95% CI 1.23-3.65) was related to the incidence of shoulder complaints. The incidence of neck complaints was increased for obesity (OR 1.81; 95% CI 1.07-3.05), work in awkward postures (OR 1.76; 95% CI 1.11-2.78) and poor/fair general health (OR 1.53; 95% CI 1.02-2.31). Recurrence of both neck and shoulder complaints was associated with chronic complaints at baseline (shoulder: OR 1.91; 95% 1.36-2.67, neck: OR 1.71; 95% 1.14-2.55) but not with work-related risk factors. **Conclusion** These results suggest that there are differences in risk factors for incidence and recurrence of neck and shoulder complaints.

1 Introduction

Shoulder and neck disorders are common problems in the general population with one-year prevalence rates ranging between 7-47 % for shoulder complaints[1,2] and 16-61% for neck complaints.[3] These complaints are characterized by high rates of incidence (18-20%) [4,5] and recovery (40-59%),[6-8] but also by high rates of recurrence (41-65%) (6-8). They may result in substantial sick leave, loss of productivity, and inability to carry out household and leisure time activities.[4,9,10]
Little is known about the factors which determine the occurrence of neck and shoulder complaints.[9-12] A few longitudinal studies reported gender (women), obesity, work in awkward postures of the neck, low social support of colleagues and high work demands as risk factors for the incidence of neck complaints.[13-16] Incident shoulder complaints were found to be related to obesity, pushing and pulling and mental distress.[17-19] Two studies[13,19] only compared the risk factors for incidence and recurrence. Eriksen et al.[13] found that high job demands and headache were risk factors for both incidence and recurrence of neck complaints, while the incidence also was associated with emotional well-being and recurrence also with gender and shoulder complaints. With respect to shoulder complaints, Miranda et al.[19] found that personal characteristics (obesity and mental distress) and work-related physical factors (physical strenuousness of work, work with hand above shoulder level, flexed trunk, twisting trunk, or rotated neck) were risk factors for incidence but only age over 45 year and overload at work were a risk factors for recurrence. Two other studies explored the course of shoulder complaints and found associations between recurrence and severity of shoulder complaints at baseline.[7,20]
The objective of this study was to assess the differences and similarities in work-related physical, psychosocial and personal risk factors for the incidence and recurrence of neck and shoulder complaints.

2 Subjects and Methods

Population Data were used from a cohort of 769 workers of nursing homes and homes for the elderly in the Netherlands. This longitudinal study with a follow-up of two years was designed to collect information on risk factors for work related musculoskeletal disorders. The cohort consisted of workers, who at inclusion had worked for at least 12 months over 10 hours a week in their job. A variety of professions were represented: nurse (n=129), care giver (n=264), office worker (n=146), kitchen worker (n=58), housekeeper and cleaner (n=49), physical therapist (n=38), maintenance worker (n=14), and various jobs (n=62). Baseline measurements, between March 1998 and March 1999, and the two follow-up measurements were performed by means of a questionnaire.

Questionnaire The baseline questionnaire on potential risk factors comprised questions about (1) personal characteristics, (2) work-related physical load, (3) work related psychosocial load and social support, (4) sports in leisure time, (5) general health status, and (6) need for recovery after a working day. A short description of these items is given.

Neck and shoulder complaints The outcomes of this study were the incidence and recurrence of shoulder and neck complaints. An incident case was defined as a subject, who had new episodes of complaints after a year free of complaints. A recurrent case was defined as a subject, who had complaints subsequent to a previous year of complaints. Additional, the prevalence was assessed by defining a prevalent case as a subject who had complaints in a given year and chronic cases were defined as subjects with complaints present almost all day with a minimal presence of at least 3 months. Chronic cases were only used as a risk factor for recurrence, not for outcome.
Data on neck and shoulder complaints were collected by means of the Nordic Questionnaire.[21] In this questionnaire a case is defined as a

person having had an episode of pain, stiffness or discomfort that was present for at least a few hours during the past 12 months. Data for neck and shoulder complaints were assessed independently of each other with questions specified for either neck or shoulder complaints. A shaded manikin accompanied the questions for both complaints (see appendix).
A prevalent case was defined as a subject, who had episodes of neck or shoulder complaints in the year before measurement. This could be either an incident case or a recurrent case.

Personal characteristics Subjects were asked about their age, gender, height, weight, level of education, and family status, using a standardized questionnaire.[22] Age was divided into four categories; (1) younger than 30, (2) 30 to 39 years, (3) 40 to 49 years, and (4) 50 to 65 years. The body mass index (BMI) was calculated by the standard formula weight/(length^2). Obesity was defined as the BMI was over 30. Family status was dichotomized in singles and people who lived together or were married.

Work related physical load Subjects were asked to fill in a 19 item questionnaire on physical load based on a modified version of the validated DMQ.[23] Seven of these items were considered relevant for this study on neck and shoulder complaints. These items were: (1) work with hands above shoulder level, (2) lifting 25 kg or more, (3) use force with arms and/or hands, (4) bending or turning the torso frequently per hour, (5) work in uncomfortable postures, (6) prolonged work in the same position of the body, and (7) repetitive movements with arms and/or hands frequently per hour. Items 2 and 3 were combined into manual material handling and items 4 and 5 were combined into work in awkward postures. A four-point scale was used with ratings ‘seldom or never’, ‘sometimes’, ‘often’ and, ‘very often’. People were supposed to be at risk if they scored ‘often’ or ‘very often’.

Work related psychological load and social support The Job Content Questionnaire[24] was used to obtain information on psychosocial aspects of work according to the demand/control hypothesis of Karasek. In this model subjects are supposedly at risk when experiencing high job demands and low job control. Job demands were measured by eleven questions on a four-point scale (never, sometimes, often, always), yielding a sum score for high work demands. The questions on work demands were related to working fast, working hard, excessive work, insufficient time to complete the work, and conflicting demands. Six questions on skill discretion and eleven questions on decision authority measured low job control. These questions concerned aspects such as required skills, task variety, learning new things, and amount of repetitive work. Workers at risk (i.e. high demands and low control) were dichotomized using the median sum scores on the job demands and job control scale. Support of colleagues and supervisor was also measured using the Job Content Questionnaire. Low support was defined as a respondent with a sum score above median.

Sports in leisure time Sports in leisure time was measured by the question: Did you exercise or participate in sport at least once a week in the past 12 months?

General health A measure of perceived general health was constructed based upon 11 dichotomized questions on general health,[25] excluding musculoskeletal symptoms. This Dutch scale has a good internal reliability (Cronbach's α = 0.86) and test-retest reliability (Pearson's r = 0.76). A sum score was calculated over all 11 items and a classification of poor/fair general health was assigned to respondents who scored in the highest half of the study population.

Need for recovery Subjects were asked to fill in eleven questions on perceived need for recovery after a working day.[26] These questions included items that focus on feelings and emotions after work such as

exhaustion, time needed to feel relaxed, energy for leisure time activities and energy for social contact with others. A sum score was calculated over all 11 items and a classification of 'need for recovery' was assigned to respondents who scored in the highest half of the study population.

3 Statistical analysis

To analyze the risk factors for incidence and recurrence of neck and shoulder complaints, Generalized Estimating Equations (GEE) was used[27,28] The GEE-analysis was performed on all subjects who completed at least one questionnaire during the follow-up period. In the GEE-model all independent variables were assumed to be time-independent, which means that only information on risk factors from the baseline measurements was used as in conventional regression analysis. These risk factors at baseline were related to shoulder complaints reported in follow-up 1 and/or follow-up 2, using two endpoints: incidence and recurrence. Thus, the analysis was stratified for subjects with shoulder complaints (or neck complaints) during the baseline survey (recurrence) and those subjects without complaints (incidence). The odds ratio expresses the association between a risk factor at baseline and the occurrence of shoulder or neck complaints during the follow-up.

The analysis was carried out with Proc Genmod in the statistical package of SAS (version 8.2). The protocol for the analysis consisted of four steps. Firstly, all independent variables were analyzed in a univariate model. Secondly, the variables with a p-value equal or less than 0.10 were included in a multivariate model by a step forward procedure. The variable with the lowest p-value was put in the model first, followed by the next lowest and so on. Covariates with a p-value lower than 0.05 remained in the model and the other variables were excluded. Age and sex were both included in the multivariable model independent of their p-value. Thirdly, we determined whether all non-

significant variables were excluded correctly by including them in the multivariate model of step two. When the model changed more than 10%, the variables was included in the multivariate model of step two. And fourthly, to keep comparable multivariate models for the incidence and recurrence, the significant risk factors in the multivariate model for incidence were added to the multivariate model of recurrence, and those for recurrence were added to the multivariate model of incidence.

4 Results

Population At baseline 769 workers agreed to participate in the study and filled in the baseline questionnaire. At 1-year follow-up 529 (68%) of the subjects filled out the questionnaire again, of whom 346 subjects (65%) responded also at 2-year follow-up. The responders and non-responders were not different according to the prevalence of neck and shoulder complaints in the year before dropout. The responders at first follow-up showed at baseline a prevalence of shoulder complaints of 38% and for the non-responders this was 34%. For the prevalence of neck complaints the corresponding figures were 36% and 40%.

Table 1

The 12-months incidence and recurrence for both neck and shoulder complaints at follow-up 1 and 2

Measurement	Neck pain		Shoulder pain	
	Incidence	Recurrence	Incidence	Recurrence
Follow-up 1	19.0%	59.0%	14.8%	63.3%
(n=529)	(n=64)	(n=113)	(n=49)	(n=126)
Follow-up 2	17.4%	64.7%	17.9%	69.4%
(n=346)	(n=40)	(n=75)	(n=42)	(n=76)

At each follow-up measurement, approximately 18 % of the workers reported incident episodes of neck or shoulder complaints, as is shown in table 1. The annual recurrence was approximately 64% for both neck and shoulder complaints. Data on the occurrence of neck and shoulder

complaints for at least one of the two follow-up measurements was available for 556 workers (71.4%). These data were used in the analysis with the GEE-method.

Table 2

Personal, work related physical and psychosocial risk factors for the 12-months incidence and the 12-months recurrence of shoulder pain; results of the univariat GEE-analyses

	Incidence (n =357)			Recurrence (n =199)		
Variable	N	Odds ratio	95% CI	n	Odds ratio	95% CI
Personal characteristics						
Gender (women)	286	1.62	0.84-3.10	187	1.29	0.74-2.27
< 30 years	60	1.00	-	34	1.00	-
30 – 40 years	100	0.95	0.47-1.93	54	0.74	0.45-1.22
40 - 50 years	122	1.72*	0.89-3.29	70	0.76	0.48-1.21
50 – 65 years	65	1.14	0.51-2.52	51	0.77	0.48-1.26
Body Mass Index > 30	37	2.23**	1.29-3.87	25	0.95	0.57-1.57
Single	81	1.12	0.69-1.85	48	0.99	0.68-1.46
Work-related physical factors						
Manual material handling	96	1.39	0.88-2.19	67	1.38**	1.00-1.92
Repetition of movement with hands or arm frequently per hour	123	1.59**	1.03-2.46	97	0.99	0.72-1.37
Work in awkward postures	79	1.44	0.89-2.31	60	1.29	0.91-1.84
Prolonged working in the same position	109	1.41	0.89-2.22	70	0.92	0.65-1.30
Working above shoulder level	32	1.22	0.60-2.49	34	1.13	0.75-1.71
Work-related psychosocial factors						
Low Job Control / High work demands	77	1.49*	0.92-2.39	66	1.21	0.86-1.69
Low Job control*	141	1.14	0.73-1.76	108	0.87	0.63-1.21
High work demands*	171	1.11	0.72-1.72	120	1.12	0.81-1.55
Low support of supervisor	153	1.27	0.82-1.96	97	0.99	0.71-1.37
Low support of colleagues	159	0.76	0.49-1.18	119	0.92	0.66-1.27
Health						
General health	136	1.35	0.87-2.09	124	1.24	0.88-1.74
Need for recovery	152	1.21	0.78-1.87	139	0.97	0.69-1.38
Duration of complaints longer than 3 months	-	-	-	52	1.72**	1.22-2.42

	Incidence (n =357)			Recurrence (n =199)		
Variable	N	Odds ratio	95% CI	n	Odds ratio	95% CI
Sport						
Exercise for at least one time a week	157	0.87	0.55-1.36	83	0.99	0.71-1.39

* Variables with a p-value > 0.05 - <0.10

** Variables with a p-value ≤ 0.05

Risk factors for incidence and recurrence of shoulder complaints

Work related physical and psychosocial factors and personal characteristics were analyzed for their associations with incidence and recurrence of shoulder complaints (Table 2). The incidence of shoulder complaints was increased for obesity (OR 2.23; 95% CI 1.29-3.87) and repetitive movement with hands or arms (OR 1.59; 95% CI 1.03-2.46). Neither health related factors nor sport activities were significantly associated with the incidence. Recurrence was associated with manual material handling (OR 1.38; 95% CI 1.00-1.92) and chronic complaints in the year before baseline (OR 1.72; 95% CI 1.22-2.42).

In the multivariate analyses (Table 3) only obese people had a higher risk on incident shoulder complaints (OR 2.12; 95% CI 1.23-3.65) and people with chronic complaints at baseline (OR 1.91; 95% CI 1.36-2.67) had a higher risk on recurrence of these complaints.

Table 3

The results of the multivariate GEE-analyses of risk factors for the 12-months incidence and recurrence of shoulder complaints

	Incidence (n=357)		Recurrence (n=199)	
Variable	Odds ratio	95% CI	Odds ratio	95% CI
BMI > 30	2.12**	1.23-3.65	0.92	0.57 - 1.51
Gender (women)	1.59	0.81-3.11	1.44	0.83 - 2.49
Younger than 30	1.00	-	1.00	-
30 – 39 years	1.04	0.51-2.10	0.78	0.36 –1.45
40 - 49 years	1.73	0.90-3.32	0.74	0.34 –1.30
50 – 65 years	1.21	0.54-2.68	0.69	0.43 – 1.12
Duration of complaints longer than 3 months	-	-	1.91**	1.36 – 2.67

** Variables with a p-value ≤ 0.05

Risk factors for incidence and recurrence of neck complaints

The relationship between work related physical, psychosocial and personal factors and neck complaints are listed in table 4. The incidence of neck complaints was increased for obese people (OR 2.21; 95% CI 1.32-3.70), work in awkward postures (OR 1.65; 95% CI 1.04-2.60), prolonged working in the same position (OR 1.61; 95% CI 1.05-2.46), manual material handling (OR 1.57; 95% CI 1.03-2.37), high job demands/low job control (OR 1.55; 95% CI 1.00-2.40), and a poor/fair general health (OR 1.62; 95% CI 1.07-2.44). The recurrence of neck complaints was related to chronic complaints at baseline (OR 1.71; 95% CI 1.17-2.47). In the multivariate analyses (table 5, page 72) obesity (OR 1.84; 95% CI 1.09-3.10), work in awkward postures (OR 1.76; 95% CI 1.11-2.78), and a poor/fair general health (OR 1.57; 95% 1.04-2.36) remained significantly associated with the incidence of neck complaints. Chronic complaints at baseline (OR 1.71; 95% CI 1.14-2.55) remained associated with the recurrence of neck complaints.

Table 4

Personal, work related physical and psychosocial risk factors for the 12-months incidence and 12-months recurrence of neck complaints; results of the univariate GEE-analyses.

	Incidence (n=352)			Recurrence (n=204)		
Variable	n	Odds ratio	95% CI	n	Odds ratio	95% CI
Personal characteristics						
Gender (women)	289	1.30	0.74-2.30	184	1.36	0.70-2.63
Younger than 30 years	59	1.00		35	1.00	-
30 – 39 years	100	1.90*	0.96-3.73	54	0.85	0.50-1.47
40 - 49 years	123	1.54	0.80-2.45	69	1.11	0.68-1.83
50 – 65 years	70	1.66	0.81-3.41	46	1.16	0.69-1.95
Body Mass Index > 30	38	2.21**	1.32-3.70	24	0.82	0.47-1.42
Single	77	1.20	0.74-1.95	52	0.81	0.53-1.24
Work-related physical factors						
Manual material handling	98	1.57**	1.03-2.37	65	1.11	0.77-1.61
Repetition of movement with hands or arm frequently per hour	125	1.45*	0.96-2.19	102	1.22	0.86-1.71

	Incidence (n=352)			Recurrence (n=204)		
Variable	n	Odds ratio	95% CI	n	Odds ratio	95% CI
Work in unpleasant postures	75	1.65**	1.04-2.60	64	1.28	0.90-1.83
Prolonged working in the same position	107	1.61**	1.05-2.46	72	1.00	0.70-1.43
Working above shoulder level	35	1.32	0.71-2.45	31	0.93	0.60-1.46
Work-related psychosocial factors						
Low job control / high work demands	87	1.55**	1.00-2.40	56	1.25	0.87-1.79
Low Job control	151	1.37	0.91-2.06	98	0.90	0.63-1.27
High work demands	178	1.22	0.81-1.83	113	0.90	0.63-1.27
Low support of supervisor	167	0.91	0.60-1.37	83	1.00	0.70-1.42
Low support of colleagues	163	1.05	0.70-1.58	115	1.16	0.81-1.66
Health						
General health	131	1.62**	1.07-2.44	129	1.18	0.82-1.71
Need for recovery	162	1.16	0.77-1.74	129	1.17	0.81-1.68
Duration of complaints longer than 3 months	-	-	-	45	1.71**	1.17-2.47
Sport						
Exercise for at least one time a week	153	0.85	0.56-1.29	87	0.91	0.64-1.29

* Variables with a p-value > 0.05 - <0.10

** Variables with a p-value ≤ 0.05

5 Discussion

The results of this longitudinal study suggest that there are differences in risk factors for the incidence and recurrence of neck and shoulder complaints. Adjusted for age and gender, obesity was related to the incidence of shoulder complaints. The incidence of neck complaints was also associated with obesity, and additionally with work in awkward postures and poor/fair general health. Recurrence of both neck and shoulder complaints, however, was only associated with chronic complaints at baseline.

The indication that risk factors may differ for incidence and recurrence was also found in other studies.[13, 19] Eriksen et al.[13] found that, for neck

complaints, emotional wellbeing was associated with incidence, while gender and shoulder complaints were associated with recurrence. They also found that high job demands and headache were associated with the 4-years incidence and recurrence of neck complaints. Miranda et al.[19] found associations with age for the one-year recurrence of shoulder complaints, while the one-year incidence was associated with age, obesity, mental stress and physical strenuousness of work.

Table 5

The results of the multivariate GEE-analyses of risk factors for the 12-months incidence and recurrence of neck complaints

		Incidence (n=352)		Recurrence (n=204)
Variable	OR	95% CI	OR	95% CI
Gender (women)	1.20	0.67-2.13	1.28	0.65-2.54
Younger than 30 years	1.00		1.00	
30 – 39 years	1.94	0.95-3.94	0.99	0.58-1.70
40-49 years	1.67	0.83-3.37	1.28	0.76-2.15
50-65 years	2.08	0.98-4.44	1.19	0.67-2.10
Body Mass Index > 30	1.84**	1.09-3.10	0.75	0.44-1.28
Work in awkward postures	1.76**	1.11-2.78	1.31	0.90-1.89
General health	1.57**	1.04-2.36	1.11	0.75-1.61
Duration of complaint longer than 3 months	-	-	1.71**	1.14-2.55

** variable with a p-value ≤0.05

In our opinion it seems defendable that specific risk factors causing the onset of shoulder and neck complaints will also play a role in initiating a recurrent episode. A possible explanation not finding similar risk factors for incidence and recurrence may be the use of self-administered questionnaires to measure physical load. Viikari-Juntura et al.[29] compared this method to task analysis and observation of physical workload. They found that subjects with complaints tend to overestimate their physical load exposure in work. To asses whether this could explain the differences we found, we compared the baseline

data of cases and symptom free subjects on physical workload stratified for the three largest occupational groups: nurses, care givers, and office workers. Within these groups, the differences in self reported physical load were reflected in the difference in the observations for back postions.[30] For different aspects of physical workload the results of this analysis suggest that subjects with complaints did not report significantly higher exposures than their healthy colleagues.

Factors related to the incidence of shoulder complaints have been studies in a few longitudinal studies.[17-19,31] Contradictory results were reported. Obesity was found to be relevant in our study and that of Miranda et al,[19] but not in the study of Leclerc et al.[31] Low job control-high job demands was risk factor in our study and that of Leclerc et al,[31] but not in the study of Miranda et al.[19] Dissimilarities in exposure definition and in the frequency of exposure may have caused these differences. Frequent repetitive movement of the hands and arms have not been evaluated in other longitudinal studies, although several cross sectional studies have found associations with repetitive work.[10,32] Working with the hands above the shoulder was not found to be significantly related to incidence in this study, although it was in others,[19,31] presumably because this exposure occurred relatively infrequently in our study. The incidence of neck complaints was associated with several physical factors (work in awkward postures, prolonged work in the same position, and manual materials handling). This finding is consistent with the results of other studies[5,11,14,33] The influence of high job demands-low job control had also been corroborated in other studies.[13.15]

Some methodological aspects possibly influenced our results. Unfortunately, we had to deal with a substantial loss to follow-up. However, it is unlikely that this had a strong influence on the outcome since we found no significant difference between the responders and the non-responders on personal characteristics and prevalence of complaints. Moreover, a recent study[34] indicated that differences in

occupational conditions and health among participants and drop out subjects did not markedly influence the risk ratios.

Three other methodological aspects are related to the case definition used by us. Firstly, we used the original case definition from the standardized Nordic Questionnaire: 'having any complaints in past year'. One might argue whether this undefined duration of complaint is relevant. Therefore, we analyzed also our data with a more strict case definition, which included complaints existing for at least 7 days in the past year as has been suggested by some authors.[16,19] The outcome of this analysis was not much different from our original findings, which can be explained by the fact that 80% of our original cases reported complaints for at least 7 days.

Secondly, the cut off point between incident and recurrent complaints was prompted by the study design with its annual measurements of the occurrence of neck and shoulder complaints. In this relatively long recall period a subject may have experienced a first spell of complaints followed by a recurrent spell in the same year. Thus, an incident case might be as well a recurrent case in terms of the number of episodes in one single year. However, we separated the recurrent cases from the incident cases by requiring the latter ones to be at least one year free of complaints. This issue of separating incident from recurrent episodes with musculoskeletal complaints has previously been noticed for low back pain.[35,36] These authors have proposed specific definitions for duration of an episode, although there is still no consensus on this topic.[35,36]

Thirdly, although shoulder and neck complaints were measured and analyzed separately, we found an overlap of 50-60% between these complaints during the three measurements. Given this overlap, it is interesting to see whether an analysis of the combination of these complaints would produce other results. Therefore, we analysed the relationship between the personal, work-related physical and psychosocial factors, and the simultaneous presence of neck and shoulder complaints. We found some differences for recurrence of

neck-shoulder complaints. Low job control/high job demands and poor/fair general health were both associated with these complaints, whereas these factors were not associated with the separate complaints. However, the observed differences may also be due to the changes in occurrence of complaints and associated sample sizes, since the number of subjects with both complaints (who are at risk for recurrence) increased and the number of healthy subjects (who are at risk for incident complaints) decreased.

In summary, we observed differences in factors related to incidence and recurrence for shoulder and neck complaints. When adjusted for age and gender, obesity was related to the incidence of shoulder complaints. The incidence of neck complaints was increased for obesity, work in awkward postures, and poor/fair general health. Recurrence of both neck and shoulder complaints was, however, only associated with chronic complaints at baseline.

6 Acknowledgement

We would like to Arianne P. Verhagen, PhD for her contributions to this chapter.

7 Reference List

1. Jacobsson L, Lindgarde F, Manthorpe R. The commonest rheumatic complaints of over six weeks' duration in a twelve-month period in a defined Swedish population. Prevalences and relationships. Scand J Rheumatol 1989;18:353-60.
2. Natvig B, Nassioy I. Musculoskeletal complaints in a population. Tidsskrift Nor Lægeforen 1994;114:323-7.
3. Nachemson AL, Waddell G, Norlund AI. Epidemiology of neck and low back pain. In: Nachemson AL, Jonsson E. Neck and Back Pain. The Scientific Evidence of Causes, Diagnosis, and Treatment. Philadelphia: Lippincott Williams & Wilkins, 2000.
4. Croft PR, Lewis M, Papageorgiou AC, Thomas E, Jayson MI, Macfarlane GJ, Silman AJ. Risk factors for neck pain: a longitudinal study in the general population. Pain 2001;93:317-25.
5. Leclerc A, Niedhammer I, Landre MF, Ozguler A, Etore P, Pietri-Taleb F. One-year predictive factors for various aspects of neck disorders. Spine 1999:24:1455-62.

6. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. BMJ 1996;313:601-2.
7. Macfarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. J Rheumatol 1998;25:1612-5.
8. Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-de Jong B. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology (Oxford) 1999;38:160-3.
9. Heijden GJ. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
10. Windt DA, Thomas E, Pope DP, Winter AF, Macfarlane GJ, Bouter LM, Silman AJ. Occupational risk factors for shoulder pain: a systematic review. Occup Environ Med 2000;57:433-42.
11. Ariëns GA, Mechelen W, Bongers PM, Bouter LM, Wal G. Physical risk factors for neck pain. Scand J Work Environ Health 2000;26:7-19.
12. Ariëns GA, Mechelen W, Bongers PM, Bouter LM, Wal G. Psychosocial risk factors for neck pain: a systematic review. Am J Ind Med 2001;39:180-93.
13. Eriksen W, Natvig B, Knardahl S, Bruusgaard D. Job characteristics as predictors of neck pain. A 4-year prospective study. J Occup Environ Med 1999;41:893-902.
14. Ariëns GA, Bongers PM, Douwes M, Miedema MC, Hoogendoorn WE, Wal G, et al. Are neck flexion, neck rotation, and sitting at work risk factors for neck pain? Results of a prospective cohort study. Occup Environ Med 2001;58:200-7.
15. Ariëns GA, Bongers PM, Hoogendoorn WE, Houtman IL, Wal G, Mechelen W. High quantitative job demands and low coworker support as risk factors for neck pain: results of a prospective cohort study. Spine 2001;26:1896-901.
16. Viikari-Juntura, E., Martikainen, R., Luukkonen, R., Mutanen, P., Takala, E. P., and Riihimäki, H. Longitudinal study on work related and individual risk factors affecting radiating neck pain. Occup Environ Med 2001;58:345-52.
17. Kaergaard A, Andersen JH. Musculoskeletal disorders of the neck and shoulders in female sewing machine operators: prevalence, incidence, and prognosis. Occup Environ Med 2000;57:528-34.
18. Hoozemans MJ, Beek AJ, Fring-Dresen MH, Woude LH, Dijk FJ. Low-back and shoulder complaints among workers with pushing and pulling tasks. Scand J Work Environ Health 2002;28:293-303.
19. Miranda H, Viikari-Juntura E, Martikainen R, Takala EP, Riihimaki H. A prospective study of work related factors and physical exercise as predictors of shoulder pain. Occup Environ Med 2001;58:528-34.
20. Van der Windt DA, Koes BW, Boeke AJ, Deville W, de Jong BA, Bouter LM. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
21. Kuorinka I, Jonsson B, Kilbom Å, Vinterberg H, Biering-Sorensen F, Andersson G, Jorgensen K. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 1987;18:233-7.
22. Central Bureau of Statistics. Netherlands health interview survey 1981-1997. SDU/publishers, The Hague, 1992.

23. Hildebrandt VH, Bongers PM, Van Dijk FJ, Kemper HC, Dul J.Dutch Musculoskeletal Questionnaire: description and basic qualities. Ergonomics 2001;44:1038-55.
24. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. J Occup Health Psychol 1998;3:322-55.
25. Van Sonsbeek JLA. The Voeg: A list of subjective health complaints. SDU/ publishers (Statistical reports M37), The Hague, 1990.
26. Sluiter JK, van der Beek AJ, Frings-Dresen MH. The influence of work characteristics on the need for recovery and experienced health: a study on coach drivers. Ergonomics 1999;42:573-83.
27. Liang K-Y, Zeger S. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13-22.
28. Twisk JW. Different statistical models to analyze epidemiological observational longitudinal data: an example from the Amsterdam Growth and Health Study. Int J Sports Med 1997;18:S216-24.
29. Viikari-Juntura E, Rauas S, Martikainen R, Kuosma E, Riihimaki H, Takala E-P, et al. Validity of self-reported physical work load in epidemiologic studies on musculoskeletal disorders. Scan J Work Environ Health 1996;22:251-9.
30. Jansen JP, Burdorf A, Steyerberg E. A novel approach for evaluating level, frequency and duration of lumbar posture simultaneously during work. Scan J Work Environ Health 2001;27:373-80.
31. Leclerc A, Chastang JF, Niedhammer I, Landre MF, Roquelaure Y. Incidence of shoulder pain in repetitive work. Occup Environ Med 2004;61: 39-44.
32. Frost P, Bonde JP, Mikkelsen S, Andersen JH, Fallentin N, Kaergaard A, et al. Risk of shoulder tendonitis in relation to shoulder loads in monotonous repetitive work. Am J Ind Med 2002;41:11-18.
33. Viikari-Juntura E, Riihimäki H, Tola S, Videman T, Mutanen P. Neck trouble in machine operating, dynamic physical work and sedentary work: a prospective study on occupational and individual risk factors. J Clin Epidemiol 1994;47:1411-22.
34. Bildt C, Alfredsson L, Punnett L, Theobald H, Torgen M, Wikman A. Effects of drop out in a longitudinal study of musculoskeletal disorders.
Occup Environ Med 2001;58:194-9.
35. Vet HC, Heymans MW, Dunn KM, Pope DP, Beek AJ, Macfarlane GJ. Episodes of low back pain: a proposal for uniform definitions to be used in research. Spine 2002;27:2409-16.
36. Wasiak R, Pransky GS, Webster BC. Methodological challenges in studying recurrence of low back pain. J Occup Rehab 2003;13:21-31.

Appendix

Preshaded manikins used in the Nordiq Questionnaire, with accompanying text as used in the questionnaire.

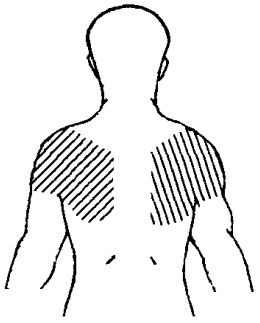


Figure 1 In this picture you can see the part of the body referred to in this questionnaire as shoulder. By shoulder problems are meant: pain, discomfort, stiffness or numbness in the shaded area. There are separate questions on neck complaints.

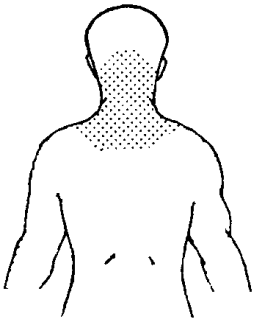


Figure 2 In this picture you can see the part of the body referred to in this questionnaire as neck. By neck problems are meant: pain, discomfort, stiffness or numbness in the shaded area, independent of adjacent areas. There are separate questions on shoulder complaints.

Chapter 5

Does this patient have instability of the shoulder or a labrum lesion?

Accepted for publication in JAMA

Abstract

Context History taking and clinical tests are commonly used to diagnose shoulder pain. Unclear is whether tests and history accurately enough to diagnose instability or intra-articular pathology (IAP). **Objective** To analyse the accuracy of clinical tests and history taking for instability or IAP. **Data sources** Relevant studies identified through searches of PubMed, EMBASE and CINAHL and bibliographies of known primary and review articles. **Study selection** Studies comparing the performance of history items or physical examination with a reference standard were included. Studies on fibromyalgia, fractures or systemic disorders were excluded. Of 1449 articles 35 were potentially eligible and 17 were selected. **Data extraction** Data were extracted on study population, clinical tests, reference tests and outcome. The studies' methodological quality (patient spectrum, verification, blinding, and replication) was assessed with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist. **Data synthesis** Six tests showed positive likelihood ratios (LRs) and confidence intervals (CIs). Tests favouring the diagnosis for establishing instability included relocation test (LR^+ 6.5; 95% CI 3.0-14.0) and anterior release test (LR^+ 8.3; 95% CI 3.6-19). Tests showing promise for establishing labral lesions included: biceps load I (LR^+ 29; 95% CI 7.3-115.0) and biceps load II tests (LR^+ 6; 95% CI 8.6-80.0), the pain provocation test of Mimori (LR^+ 7; 95% CI 1.6-32.0), and the internal rotation resistance strength test (LR^+ 25; 95% CI 8.1-76.0). The apprehension, clunk, load and shift tests, and sulcus sign proved to be less useful. Results should be cautiously interpreted because studies were completed in selected populations in orthopaedic practice, mostly assessed by the test designers and evaluated in single studies only. No accuracy studies were found for history taking or clinical tests in primary care. **Conclusion** Shoulder complaints are frequently recurrent. Instability might cause for some of these complaints. Best available evidence supports the value of relocation and anterior release

tests. Symptoms related to IAP (labral tears) remain unclear. Most promising for establishing labral tears are currently the biceps load I and II, pain provocation of Mimori and the internal rotation resistance strength tests.

1 Background

Clinical scenario A 24 year old man with a history of shoulder complaints presents to his primary care physician. At 16 years his shoulder was injured during karate. He recovered and did not notice recurrence of symptoms. At age of 21 while throwing a baseball, he developed sudden sharp left shoulder pain with a popping noise. He sensed that the arm stretched out of range. He experienced a short period with shoulder discomfort followed by recovery. Recently, he has started playing tennis and notes shoulder pain that requires cessation of play. Upon examination, the shoulder displays no swelling or atrophy. Internal and external rotation is somewhat painful, but not limited. His neck moves normally, through the full range of motion, without pain. In considering the differential diagnosis, one might wonder whether the history suggests instability of the shoulder and/or labrum lesions, and which physical examination findings confirms the diagnosis.

Why is the diagnosis important? The shoulder's wide range of motion gives us a great freedom of action, due to the shallow structure of the glenoid fossa, but lends minimal bony support for the large humeral head (see figure 1). The minimal bony support creates, however, a delicate balance between muscular and ligamentous strength.[1] Each year, 30% to 40% of adults experience shoulder discomfort causing 1% to 5 % of them to visit a general practitioner.[2-8] Although about half of the primary care patients with shoulder discomfort recover within a year, a substantial number experience continued discomfort or develop recurrent pain.[6,7,9] Instability of the glenohumeral joint, frequently combined with tears of the labrum (the cartilage rim of the glenoid), creates the continued problems for some of these patients.

Instability occurs when the shoulder's stabilizing structures provide too little control as the humerus moves on the glenoid. As a result, the upper arm fails to stay properly located in the glenoid fossa during

normal motion. Dislocation occurs when the humeral head has no attachment to the glenoid fossa, thus, a complete separation of the articular surfaces.

Subluxation is a symptomatic translation of the humeral head without complete separation.[1,10-12] The resultant symptoms and signs allow clinical classification according to the degree (dislocation or subluxation) and the direction (anterior, posterior, inferior or multidirectional) of the observed defects.[1,10-12] The incidence of shoulder dislocation is about 1.7% in the general population.[13] There are no data available in the scientific literature on the incidence or prevalence of subluxation.
Treatment of instability depends on the type and severity of the luxation detected during clinical examination, and the patient's functional deficits. The primary option, in most cases, is conservative treatment,[1,10,11] of strengthening the muscles of the shoulder and increasing the co-ordination of the shoulder girdle. The alternative is surgery, a useful treatment if the patient has recurrent dislocation without generalized ligamentous laxity or multidirectional instability.[1,10,11] Labral lesions are associated with instability, although they can occur without instability due to injuries or degeneration of the shoulder joint.[14-16] Labral lesions are classified based on their anatomical location and type of tear.[14] A frequently described labral tear is the superior labrum anterior posterior (SLAP) lesion.[14,15] The SLAP lesion is a tear is located at the superior part of the labrum that runs from the anterior to the posterior part, with or without lesions at the attachment of the long head of the biceps muscle. Surgical repairs of labral tears require an open or arthroscopic procedure.[14,15]

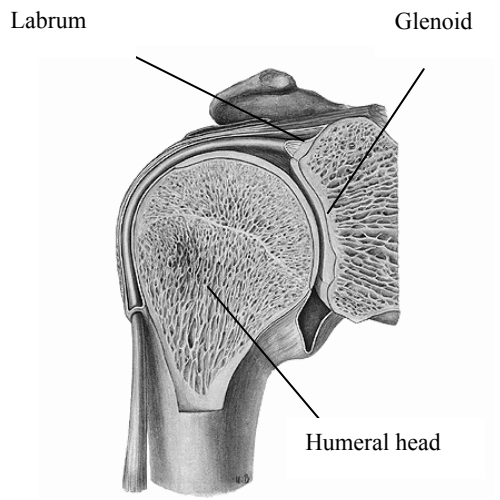
Anatomy of the shoulder The shoulder is suited for mobility. The motions of the upper arm are the result of simultaneous motions in the glenohumeral joint, the acromioclavicular joint, the sternoclavicular joint

and the scapulo-thoracic junction.[17] Shoulder instability and labral lesions affect the functioning of the glenohumeral joint.

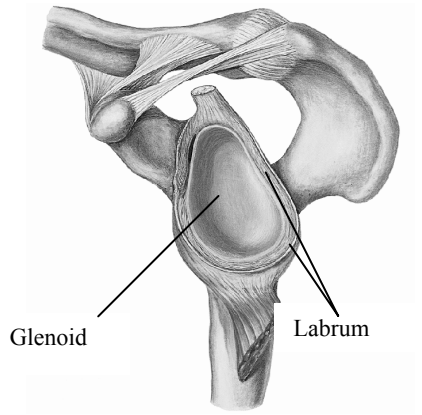
The glenohumeral joint is the articulation between the large humeral head and the small glenoid fossa of the scapula (Figure 1). The fossa is extended by the glenoid labrum (a cartilage rim) that increases the depth and surface area of the articulation.[1,14] The labrum cushions the apposition of the humeral head on the glenoid fossa, similar to the function of the menisci in the knee. A loose capsule surrounds the joint, strengthened by 3 thickenings called the anterior glenohumeral ligaments.[1]

Seventeen muscles create the movement of the shoulder.[17] The movement is a complex and subtle interaction between the 4 articulations and contributing muscles. Although knowledge of the biomechanics of the shoulder is growing, the knowledge about the relationship with clinical diagnosis is still limited. An important finding related to instability is the functioning of 4 muscles of the rotator cuff (infraspinatus, supraspinatus, teres minor and the subscapularis). These muscles play the most important roles in stabilizing the glenohumeral joint, even when the arm is in a neutral or relaxed position.[17]

Figure 1 Anatomy of the shoulder (Sobotta, 2001)



Anterior view
Cross section of the humeral head



Lateral view
(humeral head removed)

Mechanism of injuries resulting in instability or labral tears

Instability has 3 different causes. A generally known cause of anterior luxation includes a sudden traumatic fall with an outstretched arm (seen frequently in skiers) or blocked throwing movement of the arm (Figure 2). Usually, this luxation will be reduced in the field or the hospital emergency room. More typically, primary care physicians see a second type of shoulder instability created without obvious trauma and attributed to chronic gradual stretching during overhead activities in work or sport.[10] Finally, hyper laxity of the glenohumeral capsule, a less common cause of instability and often without any trauma,[1,10-12] is caused by congenital excessive joint laxity that allows the shoulder to slip in different directions (multidirectional instability). Some patients with hyper laxity of the glenohumeral capsule can dislocate their shoulder voluntarily.

The mechanisms that create labral tears without dislocation are unclear.[16] The shoulder capsule and ligaments are attached to the labrum, thus strong forces on these structures are potentially also harmful to the labrum. The occurrence of labral tears have been predominantly studied in patients with throwing injuries.[18] In this group, tears are associated with the strong forces of strain on the anterior capsule, ligaments, and labrum generated during the throwing motion. Labral tears are distinct from rotator cuff tears. A labral tear involves a tear of cartilage, while rotator cuff tear occurs in one of the tendons of the rotator cuff muscles. Instability of joint or labral tears can occur with rotator cuff injuries. However, rotator cuff injuries do not always create dislocations or labral tears. Their symptoms might be different although it is not clear from the current evidence.

Clinical presentation The diagnosis of an acute shoulder dislocation is easy to establish. It is a very painful condition and the patient will hold the arm in a fixed position.[1,10-12] However, patients with shoulder instability without dislocation present in a more subtle way. Some patients may complain about a “dead arm”-feeling.[1,10] Symptoms of pain

and functional disability seem to be non-specific for the presence of instability.[1,19] Instability of the shoulder should be considered when patients have shoulder discomfort without clear restriction of motion. A history of previous dislocation increases the likelihood of recurrent instability. Instability occurs more commonly in young people, though traumatic dislocation also occurs in older patients.[1,13]

Clinical examination of the shoulder for instability is performed to evoke recurrence of the symptoms (provocation tests) or to determine laxity of the glenohumeral joint (Table 1).[1,10] In a provocation test the humeral head is placed in a position of imminent subluxation or dislocation, which makes the patient recognize the pain-provoking movement and react with anticipated fear and/or pain (an apprehension test), see the anterior release test in Figure 2. Laxity tests of the shoulder evaluate the amount of translation of the humeral head on the glenoid in different positions of the humerus in anterior, posterior and inferior direction, such as the load and shift anterior test**.** As opposed to apprehension tests, these tests are not intended to provoke discomfort. To assess the amount of translation, specialist physicians use a classification system such as the Hawkins grading scheme (Grade 0 denotes little to no movement; grade 1 denotes when the humeral head move up onto the glenoid rim; grade 2 indicates when the humeral head can be dislocated, but spontaneously relocates; and grade 3 is when the humeral head does not relocate when the pressure is removed).[1,20] In Hawkins scheme, grades 1 to 3 are seen as a positive outcome on a laxity test

Table 1 Clinical tests for instability

Diagnostic test		
Shoulder position	Technique	Outcome
In the following tests the patient is supine		
Apprehension test		
Abducted to 90° and external rotated to 90°	Humeral head pushed in anterior direction	Pain and/or apprehension
Relocation test		
Abducted to 90° and external rotated to 90°	Humeral head pressed downwards	Relieves pain and/or apprehension

Diagnostic test Shoulder position	Technique	Outcome
Anterior release test		
Abducted to 90° and external rotated to 90°	Humeral head pressed downwards and then suddenly released	Pain and/or apprehension
In the following tests the patient is sitting or standing		
Load and shift anterior or posterior test		
Neutral position	Clinician tries to shift the humeral head in anterior or posterior direction with one hand, and stabilises the patient with the other	Does not evoke discomfort. Degree of anterior or posterior laxity is evaluated by Hawkins grading scheme
Sulcus sign		
Neutral	Arm is pulled vertically downwards	Positive when sulcus become visible between acromion and humeral head

When laxity is present in more than one direction, the diagnosis of multidirectional instability is considered and the patient should be examined for generalized ligamentous laxity (laxity in more joints of the body).[1,10-12] There are no uniformly accepted clinical criteria for generalized ligamentous laxity. One might suspect this type of laxity when finding positive laxity tests in both shoulders. Other examples of hyperlaxity include the ability to hyperextend the elbows and a positive thumb-to-forearm test whereby the patient can pull his or her thumb back to the point of touching the forearm. Typically, such patients will know that they can perform “joint tricks” that demonstrate their “loose joints.”

Table 2 Clinical tests for Labral Tears

Diagnostic test Shoulder position	Technique	Outcome
In the following test the patient is supine		
Biceps load I (-II)‖ test		
Abducted to 90° (-120°-) and fully external rotated, elbow flexed to 90°	Clinician applies force to extend the elbow as patient resist	Pain
Compression rotation		
Abducted to 90° and external	Axial load place on shoulder	Pain or clicking

Diagnostic test Shoulder position	Technique	Outcome
rotated to 90°, elbow flexed to 90°,	while rotated and circumducted (similar to Mcmurray knee test)	
In the following tests the patient is sitting or standing		
Active Compression (O'Brien)		
Forward flexed to 90°, abducted to 10-15°, fully internal rotated.	Clinician stands in front of the patient and pushes the upper arm down as the patient resist. Repeated with shoulder in fully external rotation	Pain in the first manoeuvre, reduced or eliminated in the second
Test of Speed		
Forward flexed to 90°, elbow fully extended, forearm fully supinated	Downward force applied to the forearm	Pain in anterior shoulder
Tenderness of bicipitial groove		
Neutral	Palpating the bicipital groove	Pain
Test of Yergason		
Neutral with elbow flexed to 90°	Patient supinates forearm against force applied by clinician, who simultaneously palpates the biceps tendon	Pain in the biceps tendon
Test of Mimori		
Abducted to 90° and external rotated to 90°, elbow flexed to 90°, forearm supinated	Forearm is pronated	Pain
SLAP-prehension		
Forward flexed to 90°	Arm is rotated internally	Pain or clicking
Test of Zaslav (internal rotation strength)		
Abducted to 90° and external rotated to 80°	Patient resist external rotaton force applied by the clinician, followed by internal applied force	Good strength in external rotation and apparent weakness in internal rotation

‖ The biceps load II is performed similar to the biceps load I test, the only difference is the 120° abduction of the shoulder

Patients with labral tears present with a variety of symptoms.[16] Snyder[14] suggested that the most common clinical symptoms are deep shoulder pain, pain with overhead activities, or painful catching, popping or clicking. Stetson and Templin[21] suggested that these symptoms were not specific for labral tears since they mimic the presence of impingement disorders, rotator cuff tears or other shoulder problems. Although an obvious clinical presentation for labral tears cannot be described, clinicians should consider the diagnosis when the shoulder

pain is related to a traumatic injury that involved substantial forces on the glenohumeral joint (e.g., falling while skiing).
Clinical tests for detecting labral tears (table 2) provoke symptoms by compressing the humerus into the glenoid in an attempt to catch the labral fragment between the bony structures (compression rotation test).[22] Another eponymous test to evoke symptoms by rotating the humerus passively or actively is Mimori's test,[18] shown in figure 2. Alternative physical examination maneuvers reproduce shoulder symptoms by asking the patient to resist the
Signs and symptoms for shoulder instability and intra-articular pathology (labral tears) have to be accurate in order to add appropriate diagnostic information. We reviewed the literature on the accuracy of diagnostic studies for shoulder instability and intra-articular pathology.
force of the clinician while the arm is held in a fixed position, such as the biceps load II test[23] shown in figure 2.

2 Methods

This review is based on the guidelines for systematic reviews of studies evaluating the accuracy of diagnostic tests[24] identified through the PubMed (1966-2003), EMBASE (1980-2001) and CINAHL (1982-2001) databases. To retrieve all relevant publications related to diagnosing shoulder complaints in adults, the term *exp shoulder* was searched. In addition, text word searches were completed for *glenohumeral, scapula, clavicula, acromion, rotator cuff, supraspinatus, supra-spinatus, infraspinatus, infra-spinatus, serratus anterior,* and *subscapularis.* Diagnostic studies were retrieved by exploding *sensitivity and specificity*, with additional textword searches of *specificity, false negative, screening and accuracy* based on the search strategy of Deville et al. [25] In addition, bibliographies of known primary and review articles were also examined. One reviewer (JJL) screened abstracts of the retrieved citations on: clinical tests, sensitivity and specificity figures,

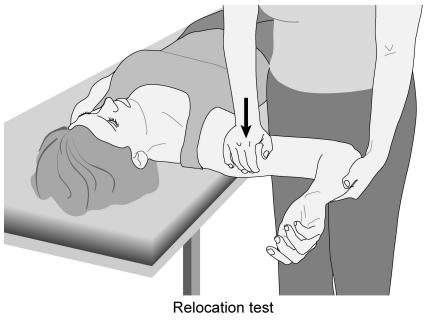
and shoulder pain. Relevant articles were obtained from the library and their reference lists were screened to find additional studies.

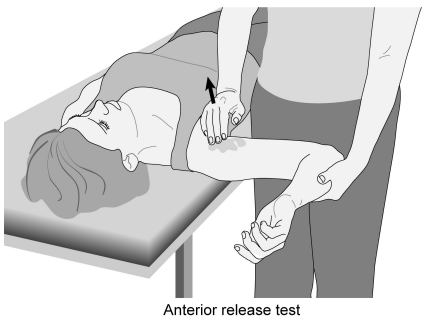
Studies were screened by 2 reviewers (JJL, BWK) and had to meet the following inclusion criteria: (1) description of clinical tests for instability or intra-articular pathology (IAP) of the shoulder, (2) use of a reference (gold) standard, (3) specification of sensitivity and specificity and (4) publication in English, Dutch or German. Studies were excluded if the diagnoses included fibromyalgia, or systemic disorders such as rheumatoid arthritis, fractures, tumours or strokes. We selected studies that compared a clinical test to surgical or arthroscopic findings, not to non-invasive imaging tests (e.g., magnetic resonance imaging, ultrasonography, or computer assisted tomography).

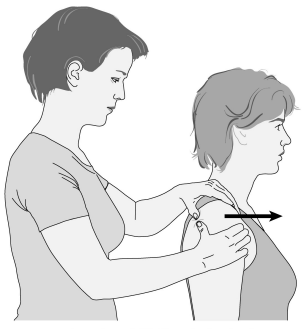
While these imaging tests may be useful confirming the presence of instability or IAP they have a sensitivity of only 60-90% depending on the type of injury and in comparison to surgery or arthroscopy.[26] Approximately 10 to 20% of patients with a normal reading on Magnetic Resonance Imaging or ultrasonography[26] may still have shoulder instability or labral tears. Thus, these non-invasive tests might ultimately prove useful as a pragmatic reference standard for some physicians, although the presence of verification bias (no surgery or arthroscopy implemented when the non-invasive study is normal) and possible low sensitivity creates uncertainty when reviewing the utility of the clinical examination.

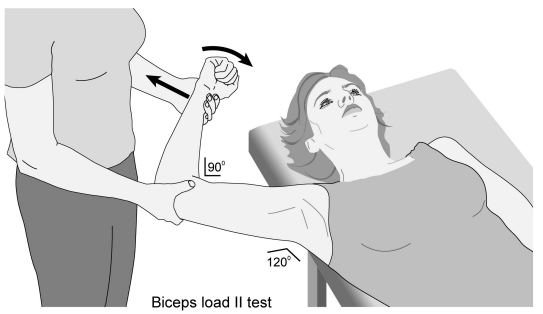
For each study, details were extracted on study population (setting, sampling, age, sex and diagnosis), clinical tests, reference tests and outcome (sensitivity and specificity). When raw data were available, likelihood ratios were calculated for individual findings, thereby describing the increase in odds that the patient had shoulder instability when a symptom or sign was present or the opposite effect when a sign of symptom was absent.

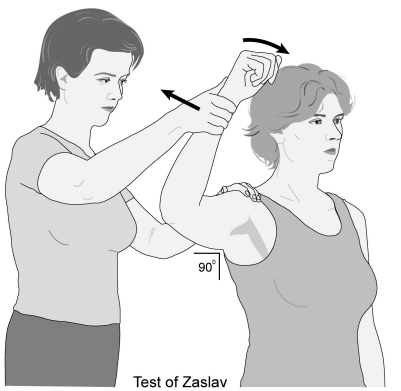
Figure 2 Examples of clinical shoulder tests; relocation test, anterior release test and load and shift anterior test for diagnosing instability and biceps load II test, Mimori´s test and the Test of Zaslav for establishing labral tears. Test techniques are described in Table 2.











The methodological quality of the studies was evaluated by 2 reviewers (APV, JJL) using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS).[27] This list includes 14 questions about the spectrum of patients studied, selection criteria, test verification, test description, blinding, uninterpretable results and study withdrawals. These questions could be scored as positive if the item was fulfilled, negative if the item was not fulfilled, or unclear if the item was not described. The limitations of each study were described. The studies were not allocated into arbitrary categories of low, medium or high quality.

3 Results

Our search strategy used a broad spectrum of terms for the shoulder, yielding about 21,000 articles. Combined with the search strategy of Deville et al.[25] on diagnosis this resulted in 1449 abstracts from the three databases. About 130 abstracts contained information on shoulder disorders and diagnostic outcome measurements. However, most of the articles evaluated sonography versus surgery, magnetic resonance imaging versus surgery, or one type of magnetic resonance imaging versus another type.

Formal reviews were conducted for 35 articles that evaluated clinical tests. Seventeen studies[16,18,19,21-23,28-38] met the selection criteria for inclusion in this review (Table 3). Eighteen studies were excluded: 11 because no information on instability or intra-articular pathology was presented,[39-49] 4 because data were missing on sensitivity and specificity or clinical tests,[50-53] and 3 because they were published in French.[54-56] Of the 17 studies that were selected: 5 enrolled patients when the clinician suspected shoulder instability; [19,33,35,37,38] and 12 enrolled patients when the clinician suspected labral tears or other IAP. All studies were conducted in orthopaedics clinics.

Each study evaluated a varying number of clinical tests, but lacked data on history. Surgery was used as a reference test in 6

studies[19,29,30,33,35,37] , and arthroscopy in 11[16,18,21-23,28,31,32,34,36,38]. Five clinical tests (apprehension test,[19,38] relocation test,[19,38] active compression test,[21,29] anterior slide test[22,34] and the test of Speed[30,38]) were evaluated in more than one study. Two studies reported the clinical examination of the shoulder under anaesthesia using the same protocol.[33,37] These studies were not pooled due to lack of clinical homogeneity in study populations. Although most studies had the same inclusion criterion for participant selection: 'having a surgery or arthroscopy for shoulder complaints', the selection standards for undergoing surgery or arthroscopy were unclear. Hence, the constitution of the population might have differed. In addition, different end points of the diagnoses made it impossible to evaluate the impact of the diagnostic threshold for sensitivity and specificity.

Table 3 Study Characteristics

Study	Selection criteria	Sample (n, mean age, % women)‡	Index test	Limitations§
Retrospective design; Arthroscopy as reference test				
Berg et al.[37] 1998	Identified SLAP lesions during arthroscopy	n = 66 - -	SLAP-prehensiontest	a c e g h l
Prospective design; Arthroscopy as reference test				
Guanche et al. [16] 2003	First arthroscopy for shoulder pain, complete range of motion under anesthesia	n = 61 38 years 19%	Active compression test Anterior apprehension test Crank test Relocation test Test of Speed Test of Yergason Tenderness in bicipital groove	a f g
Kibler[35] 1995	Isolated glenoid labral tear or, partial thickness rotator cuff pathology or, bankart lesion or, capsular deficiency or,	n = 226 - 33%	Anterior slide test	a b c d e g h i j

Study	Selection criteria	Sample (n, mean age, % women)‡	Index test	Limitations§
	25° internal rotation deficit			
Kim et al.[27] 1999	Arthroscopy for unilateral recurrent anterior shoulder dislocation (based on physical examination, plain X-ray and MRI) with a Bankart lesion Exclusion: multidirectional instability	n = 75 25 years 15%	Biceps load test I	a b c f g h
Kim et al.[28] 2001	Arthroscopy for shoulder problems Exclusion: dislocation; stiff shoulder	n =127 31 years 30%	Biceps load test II	a b c f g
Liu et al. [23] 1996	Shoulder surgery after failure of conservative treatment Exclusion: traumatic dislocation; weakness of m.subscapularis	n = 62 28 years 22%	Crank test	a c e f g
McFarland et al. [22] 2002	Diagnostic arthroscopy for shoulder pain	n =426/ 604[#] - -	Compression rotation test Anterior slide test Active compression test	a b g
Mimori et al.[18] 1999	Shoulder pain during throwing motions Exclusion: instability; indications of rotator cuff tears on MRI or arthrography	n = 32 21 years 6%	Crank test Anterior apprehension test in external and internal rotation	a b c d g h
Stetson et al. [21] 2002	Diagnostic arthroscopy after failure of conservative treatment	n = 65 46 years 31%	Crank test Active compression test	a b c g h j
T'Jonck et al. [39] 2001	Shoulder arthroscopy due to disabling shoulder pain Exclusion: > 65 years; previous surgery of shoulder;	n = 71 - 45%	Active compression test Apprehension test Clunk test Lift off test Load and shift test	a b g

Study	Selection criteria	Sample (n, mean age, % women)‡	Index test	Limitations§
	interaction with complaints in elbow or neck		Posterior stress test Release test Relocation test Resistance test external rotation Test of Speed Sulcus sign	
Prospective design; Surgery as reference test				
Bennett et al. [25] 1998	Surgery for shoulder pain	n = 45 - 31%	Test of Speed	a b c g
Cofield et al. [32] 1993	Surgery after referral for suspected recurrent instability	n = 55 29 years 27%	Laxity tests under anesthesia in anterior, posterior, inferior, anterior-inferior and posterior-inferior direction	a b c f g
Gross et al. [36] 1997	Subluxation or gross dislocation on examination under anesthesia, abnormal excursion during arthroscopic examination, Hill Sachs lesion or Bankart lesion	n = 82/ 100‖ 37 years 38%	Anterior release test	a b c f g h
O'Brien et al.[24] 1998	Shoulder pain	n = 268¶ - -	Active compression test	a b c d e f g h i j
Oliashirazi et al. [38] 1999	Shoulder surgery for unilateral traumatic recurrent anterior instability	n = 30 23 years 17%	Laxity tests under anesthesia in anterior, posterior, inferior, anterior-inferior and posterior-inferior direction	a b f g h
Speer et al. [19] 1994	Shoulder surgery; subtle anterior instability. Exclusion: Treatable/ observable rotator cuff lesions; multidirectional	n = 100 - -	90°/ 90° relocation test	a b f g

Study	Selection criteria	Sample (n, mean age, % women)‡	Index test	Limitations§
	instability			
Zaslav [26] 2001	Shoulder surgery after failure of conservative treatment; positive Neer overhead sign	n = 110/ 115** 44 years 41%	Internal rotation resistance strength test	a c g

‡ If data was not given for age or % females, this was indicated with '-'; § Key to limitations: a. Spectrum bias possible: patient on the list for surgery or arthroscopy; b. Selection criteria for waiting list entry not described; c. Disease progression bias possible: time between index and reference test not described; d. Partial verification bias; part of the sample did not receive the reference test; e. Incorporation bias: results of index test are used to establish the final diagnosis. f. The execution of the reference test was not described, which causes problems with the replication of the study. g. Blinding unclear: the reference test might have been interpreted with knowledge of the index test or visa versa; h. Unclear if same clinical data (radiography, MRI or other diagnostic information) would be available in daily practice i. Unclear if uninterpretable or intermediate test results were reported. j. Unclear if all patients who entered the study were accounted for (withdrawals). ‖ 18 patients retrospectively excluded for dual diagnosis; ¶ If only a part of the study population received the reference standard, this leaded to verification bias. This type of bias will influence the test performance, it might over- or underestimate the overall diagnostic accuracy. In this study it is likely to give an overestimation of the sensitivity because only the most sever cases have had the reference test and were a priori more likely to have the target disorder. # 178 patients retrospectively excluded for various reasons ** Five patients removed based on physical findings

Accuracy of signs and symptoms related to instability and labral tears No diagnostic studies assess the value of history taking in diagnosing instability. Four provocation tests for instability are presented in Table 4. The relocation test[38] and the anterior release test[35] have the best properties for increasing the likelihood of instability (relocation test[38] LR^+ 6.5 (95% CI 3.0-14.0) and LR^- 0.18 (95% CI 0.07-0.45); anterior release test[35] LR^+ 8.3 (95% CI 3.6-19) and LR^- 0.09 (95% CI 0.03-0.27)). The relocation test does not work as well in determining more subtle degrees of anterior instability as opposed to more obvious cases of instability, although we were unable to evaluate the CI around the LRs for detecting less significant instability.[19] The apprehension test and the clunk test were both of limited value, due to low specificity and low sensitivity, respectively.

Establishment of instability was not confirmed or ruled out with the sulcus sign[38] or the load and shift anterior posterior laxity tests.[38] The likelihood of instability increased when laxity tests were performed

under anaesthesia (LR^+ 13; 95% CI: 3.9-43),[33] however these tests cannot be performed in the general medical practice (due to the use of anaesthesia).

Table 4 Diagnostic accuracy of physical examination for instability of the shoulder

Study	Diagnosis	No. of Shoulders	Sensitivity (raw data)*	Specificity	LR^+ [95% CI]	LR^-
Apprehension test						
38	Instability	72	0.88 (23/26)	0.50 (23/46)	1.8 [1.3-2.5]	0.23 [0.08-0.69]
19	Subtle anterior instability	100	*Pain* 0.54	0.44		
19			*Apprehension* 0.68	1.00		
Relocation test						
39	Instability	72	0.85 (22/26)	0.87 (40/46)	6.5 [3.0-14.0]	0.18 [0.07-0.45]
19	Subtle anterior instability	100	*Pain* 0.30	0.58		
19			*Apprehension* 0.57	1.00		
Clunk test						
38	Instability	72	0.35 (9/26)	0.98 (45/46)	16 [2.1-119]	0.67 [0.50-0.89]
Anterior release test						
38	Instability	72	0.85	0.87		
35	Occult instability	100	0.92 (34/37)	0.89 (40/45)	8.3 [3.6-19]	0.09 [0.03-0.27]
Laxity tests						
Load and shift posterior test						
38	Instability	72	0 (0/26)	1.00 (46/46)	1.7 [0.0-83.0]	0.99 [0.93-1.1]
Sulcus sign						
38	Instability	72	0.31 (8/26)	0.89 (41/46)	2.8 [1.0-7.7	0.78 [0.59-1.00]
Load and shift anterior test						
38	Instability	72	0.54 (14/26)	0.78 (36/46)	2.5 [1.3-4.8]	0.59 [0.38-0.92]
Examination under anesthesia						
33	Instability	55	1.00 (25/25)	0.93 (28/30)	13 [3.9-43]	0.02 [0-0.31]

Study	Diagnosis	No. of Shoulders	Sensitivity (raw data)*	Specificity (raw data)*	LR^+ [95% CI]	LR^- [95% CI]
37	Anterior instability	60	0.83 (25/30)	1.00 (30/30) †	51 [3.2-801]	0.18 [0.08-0.38]

* If data of the two by two table was presented in the study the numbers for calculation of sensitivity and specificity figures are given between brackets; † The healthy contra lateral shoulders of the subjects (n=30) were used as control. Hence, the specificity value and likelihood ratios have been presumably overestimated

The possibility of detecting labral tears by arthroscopy has renewed interest in clinical tests for detecting affected patients. Thirteen studies[16,18,21-23,28-32,34-36] have evaluated 14 clinical signs, and for 8 of these [18,21-23,28,29,32,34] allowed calculation of positive and negative LRs (Table 5). The anterior slide test,[22,34] the crank test[16,21,28] and the active compression test[16,21,22,29] were promising when their designers evaluated them. However, the accuracy and LRs found by other researchers were far less hopeful. Therefore, optimism should be reserved for test results that have not been duplicated in subsequent studies. The biceps load I[32] (LR^+ 29; 95% CI: 7.3-115.0), the biceps load II test[23] (LR^+ 26; 95% CI: 8.6-80.0), the pain provocation test of Mimori[18] (LR^+ 7.2; 95% CI: 1.6-32.0), and the internal rotation resistance strength test[31] (LR^+ 25; 95% CI: 8.1-76.0) need confirmation before they become widely adopted. Conflicting evidence was found for the test of Speed.[16,30] In general, most of the evaluated clinical signs appear to have a high specificity that leads to high positive LRs. A few tests also have an excellent sensitivity that, if confirmed, would make them useful for ruling out labral tears without arthroscopy.

Table 5 Diagnostic accuracy of physical examination for labral tears

Study	Diagnosis	No. of shoulders	Sensitivity (raw data)*	Specificity (raw data)*	LR^+ [95% CI]	LR^- [95% CI]
<u>Anterior apprehension test</u>						
16	Labral tear (including SLAP)	60	0.40	0.87		
16	SLAP lesion only	60	0.30	0.63		
<u>Active compression test (O'Brien test)</u>						
21	Labral tear	65	0.54 (14/26)	0.31 (12/39)	0.8 [0.5-1.2]	1.5 [0.8-2.8]

Study	Diagnosis	No. of shoulders	Sensitivity (raw data)*	Specificity	LR^+ [95% CI]	LR^-
29	Labral tear	206	1.00 (53/53)	0.98 (150/153)	21.0 [10.0-42.0]	0.01 [0.0-0.16]
29	Acromial joint pathology	212	1.00 (55/55)	0.96 (150/157)	44 [16-123]	0.01 [0.0-0.16]
22	SLAP lesion	409†	0.47 (18/38)	0.55 (203/371)	1.0 [0.7-1.4]	0.96 [0.70-1.30]
16	Labral tear (including SLAP)	60	0.63	0.73		
	SLAP lesion only	60	0.54	0.47		
Anterior slide test						
34	Superior glenoid labral tear	226	0.78 (69/88)	0.92‡ (125/138)	8.3 [4.9-14.0]	0.24 [0.16-0.36]
22	SLAP lesion	419†	0.07 (3/38)	0.83 (62/381)	0.5 [0.2-1.5]	0.99 [1.10-1.20]
Biceps load I test						
32	SLAP lesion	74	0.83 (10/12)	0.98 (62/63)	29.0 [7.3-115.0]	0.09 [0.01-0.58]
Biceps load II test						
23	SLAP lesion	127	0.90 (35/38)	0.96 (85/89)	26.0 [8.6-80.0]	0.11 [0.04-0.28]
Compression rotation test						
22	SLAP lesion	303 †	0.24 (7/29)	0.76 (207/274)	1.0 [0.5-2.0]	1.00 [0.81-2.10]
Crank test						
28	Labral tears	62	0.91 (29/32)	0.93 (28/30)	14.0 [3.5-52.0]	0.10 [0.03-0.29]
21	Labral tears	65	0.46 (12/26)	0.56 (22/39)	1.1 [0.6-1.9]	0.95 [0.61-1.50]
16	Labral tears (including SLAP)	60	0.40	0.73		
16	SLAP lesion only	60	0.39	0.67		
Internal rotation resistance strength test						
31	Internal arti-cular deran-gement	110	0.88 (23/26)	0.96 (81/84)	25.0 [8.1-76.0]	0.12 [0.04-0.35]
Pain provocation test of Mimori						
18	Superior Labral tears	32	1.00 (22/22)	0.90 (9/10)	7.2 [1.6-32.0]	0.03 [0.00-0.47]
Relocation test						
16	Labral tears (including SLAP)	60	0.44	0.87		
	SLAP lesion only	60	0.36	0.63		
SLAP-prehension test						

Study	Diagnosis	No. of shoulders	Sensitivity (raw data)*	Specificity	LR^+ [95% CI]	LR^-
36	SLAP lesion	66	0.82 (54/66)	-		
Tenderness of bicipital groove						
16	Labral tears (including SLAP)	60	0.44	0.40		
16	SLAP lesion only	60	0.48	0.52		
Test of Speed						
30	Biceps pathology (including labral lesion)	46	0.90 (9/10)	0.14 (5/36)	1.1 [0.8-1.3]	0.72 [0.10-5.50]
16	Labral tears (including SLAP)	60	0.18	0.87		
16	SLAP lesion only	60	0.09	0.74		
Test of Yergason						
16	Labral tear (including SLAP)	60	0.09	0.93		
16	SLAP lesion only	60	0.12	0.96		

* If data of the two by two table was presented in the study the numbers for calculation of sensitivity and specificity figures are given between brackets.; † The authors stated in their article; "Patient number for each test were not equal because the test were published at different times (namely, the compression rotation test in 1990, the anterior slide test in 1995 and the active compression test in 1998)."; ‡ Healthy subject were included, therefore the specificity value was presumably overestimated.

Limitation of the literature The results of the presented studies pose some limitations and should be interpreted with caution. The limitations for each study are presented in Table 3. The diagnostic studies were all executed in specialized care; therefore, the optimal spectrum of disease was defined as patients visiting an orthopaedics clinic with shoulder pain. However, in 15 studies[16,19,21-23,28,30-38] patients were selected from waiting lists for shoulder surgery or shoulder arthroscopy. In these studies, spectrum bias cannot be excluded. Besides, this selection criterion resulted in a highly selected group of patients with severe shoulder disorders, which is also noticeable in the high prevalence values (15-100%) of instability and labral lesions. A high prevalence among study subjects reduces the opportunity to detect both false-positive and true-negative results, which will overestimate the sensitivity

and underestimate the specificity when the test is applied to patient populations with a lower prevalence of disease. It is likely that clinical findings in daily medical practice have lower sensitivity but higher specificity than suggested in the available literature.
Other limitations of the existing literature include modest sample sizes and methodological problems. Twelve[18,19,21-23,29,30,32,33,35,36,38] out of the 18 studies did not describe the procedure for patient selection. The time between index and reference test was unknown in 11 studies[18,23,28-36] and the details of the reference test were missing or unclear in 9 studies.[16,19,23,28,29,32,33,35,37] Furthermore, in 16 studies[16,18,19,21-23,28-36,38] it was unclear whether the examiner of the reference test was blinded for the index test; in one study it was evident that the examiner was not blinded.[37] These methodological problems complicate reproduction of the study results and may have possibly biased the outcome.

Resolution of the clinical scenario Primary care physicians may consider the diagnosis of instability with or without a labral tear for this 24-year-old. The history of trauma at a young age and recurrent shoulder problems associated with a symptom that might have represented an acute dislocation (pop with an excessive stretch), means that the attending physician may consider clinical tests to assess for instability and labral tears, but diagnostic accuracy would still be uncertain. Since the patient might opt for surgical repair, primary care physician might consult an orthopaedist to confirm the diagnosis and optimal management strategies for this patient's case.

4 Bottom line

The available evidence suggests that the relocation test and the anterior release test are best for establishing diagnosis of instability. For labral tears the biceps load I and II tests, the pain provocation test of

Mimori and the internal rotation resistance strength test have the best diagnostic performance characteristics (tests are shown in Figure 2). However, these results are based on single studies done in groups of selected patients who were evaluated by specialists. Despite the high prevalence of shoulder disorders in the general population, we are uncertain whether the diagnostic value of these tests, or combinations of tests, will be similar when used in primary care. Nonetheless, an understanding of the tests used in a specialist practice gives primary care physicians the opportunity to focus on physical examination manoeuvres that might improve diagnostic skills. Although we recommend that clinicians take a careful history of the mechanism of shoulder injury, the role of the patient's history for diagnosing the presence of instability or labral tears has not been studied. A comparison of relevant historical characteristics of patients with shoulder complaints, physical examination findings, and non-invasive images (e.g. magnetic resonance imaging) along with arthroscopy or surgical results would greatly enhance the knowledge base of primary care physicians who first to evaluate shoulder conditions.

5 Acknowledgement

We would like to thank David Simel, MD MHs for his critical comments on the manuscript.

6 Reference List

1. Matsen FA, Thomas SC, Rockwood CA. Anterior glenohumeral instability. In: Rockwood CA, Matsen FA ed . *The shoulder, Volume 1*. 1990 Philadelphia: W.B. Saunders Company; 526-622.
2. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: the influence of case definition**.** *Ann Rheum Dis* 1997;56:308-312.
3. Urwin M, Symmons D, Allison T, et al. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Ann Rheum Dis* 1998;57:649-655.

4. Makela M, Heliovaara M, Sainio P, et al. Shoulder joint impairment among Finns aged 30 years or over: prevalence, risk factors and co-morbidity. *Rheum*atology 1999;38:656-662.
5. Natvig B, Nassioy I. Musculoskeletal complaints in a population. Occurence . *Tidsskrift for Den Norske Laegeforening* 1994;114:323-327.
6. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. *BMJ* 1996;313:601-602.
7. Van der Windt DA, Koes BW, de Jong BA. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis* 1995;54:959-964.
8. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 2003;102:167-178.
9. Macfarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. *J Rheum* 1998;25:1612-1615.
10. Pollock RG, Flatow EL. Classification and Evaluation. In: Bigliami LU ed. The Unstable shoulder. 1996, Rosemont:Am Ac Orthop Surg; 25-36.
11. Warner JJ. Overview: Avoiding Pitfalls and managing complications and failures of instability surgery. In: Warner JJ, Iannotti JP, Gerber C. Complex and Revision Problems in Shoulder Surgery. 1997, Philadelphia;Lippincott-Raven Publishers: 3-8.
12. Gerber C. Observations on the classification of instability. In: Warner JJ, Iannotti JP, Gerber C. Complex and Revision Problems in Shoulder Surgery. 1997, Philadelphia;Lippincott-Raven Publishers:9-18.
13. Hovelius L. Incidence of shoulder dislocation in Sweden. *Clin Ortho Rel Res* 1982;127-131.
14. Snyder SJ. Labral lesions (Non_Instability) and SLAP lesions. In: Snyder SJ. Shoulder arthroscopy. 1994, New York; McGraw-Hill:115-131.
15. Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ. SLAP lesions of the shoulder. *Arthroscopy* 1990;6:274-279.
16. Guanche CA,Jones DC. Clinical testing for tears of the glenoid labrum. *Arthroscopy* 2003;19:517-523.
17. Van der Helm FC. A finite element musculoskeletal model of the shoulder mechanism. *Journal of Biomechanics* 1994;27:551-569.
18. Mimori K, Muneta T, Nakagawa T, Shinomiya K. A new pain provocation test for superior labral tears of the shoulder. *Am J Sports Med* 1999;27:137-142.
19. Speer KP, Hannafin JA, Altchek DW, Warren RF. An evaluation of the shoulder relocation test. *Am J Sports Med*. 1994;22:177-183.
20. McFarland EG, Torpey BM, Curl LA. Evaluation of shoulder laxity.Sports Med. 1996,22:264-272
21. Stetson WB, Templin K. The Crank test, the O'Brien test, and routine magnetic resonance imaging scans in the diagnosis of labral tears. *Am J Sport Med* 2002;30:806-809.

22. McFarland EG, Kim TK, Savino RM. Clinical assessment of three common tests for superior labral anterior-posterior lesions. *Am J Sports Med* 2002;30:810-815.
23. Kim SH, Ha KI, Ahn JH, Kim SH, Choi HJ. Biceps load test II: A clinical test for SLAP lesions of the shoulder. *Arthroscopy* 2001;17:160-164.
24. Deville WL, Buntinx F, Bouter LM, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002;2:9
25. Deville WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;53:65-69.
26. Dinnes J, Loveman E, McIntyre L, Waugh N. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review.Health Technol Assess. 2003;7:1 -166
27. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
28. Liu SH, Henry MH, Nuccion SL. A prospective evaluation of a new physical examination in predicting glenoid labral tears. *Am J Sports Med* 1996;24:721-725.
29. O'Brien SJ, Pagnani MJ, Fealy S, McGlynn SR, Wilson JB. The active compression test: a new and effective test for diagnosing labral tears and acromioclavicular joint abnormality. *Am J Sports Med* 1998;26:610-613.
30. Bennett WF. Specificity of the Speed's test: arthroscopic technique for evaluating the biceps tendon at the level of the bicipital groove. *Arthroscopy*. 1998;14:789-796.
31. Zaslav KR. Internal rotation resistance strength test: a new diagnostic test to differentiate intra-articular pathology from outlet (Neer) impingement syndrome in the shoulder. *J Shoulder Elbow Surg*. 2001;10:23-27.
32. Kim SH, Ha KI, Han KY. Biceps load test: a clinical test for superior labrum anterior and posterior lesions in shoulders with recurrent anterior dislocations. *Am J Sports Med* 1999;27:300-303.
33. Cofield RH, Nessler JP, Weinstabl R. Diagnosis of shoulder instability by examination under anesthesia. *Clin Orthop* 1993;291:45-53.
34. Kibler WB. Specificity and sensitivity of the anterior slide test in throwing athletes with superior glenoid labral tears. *Arthroscopy* 1995;11:296-300.
35. Gross ML,Distefano MC. Anterior release test. A new test for occult shoulder instability. *Clin Orthop* 1997;339:105-108.
36. Berg EE, Ciullo JV. A clinical test for superior glenoid labral or 'SLAP' lesions. *Clin J Sport Med* 1998;8:121-123.
37. Oliashirazi A, Mansat P, Cofield RH, Rowland CM. Examination under anesthesia for evaluation of anterior shoulder instability. *Am J of Sports Med* 1999;27:464-468.
38. T'Jonck L, Staes F, Smet L, Lysens R. The relationship between clinical shoulder tests and the findings in arthroscopic examination. *Geneeskunde en Sport* 2001;34:15-24.

39. Lyons AR, Tomlinson JE. Clinical diagnosis of tears of the rotator cuff. *J Bone Joint Surg Br* 1992;74:414-415.
40. Leroux JL, Thomas E, Bonnel F, Blotman F. Diagnostic value of clinical tests for shoulder impingement syndrome. *Rev Rhum Engl Ed* 1995;62:423-428.
41. Ure BM, Tiling T, Kirchner R, Rixen D. Zuverlassigkeit der klinischen Untersuchung der Schulter im Vergleich zur Arthroskopie. Eine prospektive Studie. *Unfallchirurg* 1993;96:382-386.
42. Hertel R, Ballmer FT, Lombert SM, Gerber C. Lag signs in the diagnosis of rotator cuff rupture. J Shoulder Elbow Surg 1996;5:307-313.
43. Litaker D, Pioro M, El Bilbeisi H, Brems J. Returning to the bedside: using the history and physical examination to identify rotator cuff tears. *J Am Ger Soc* 2000;48:1633-1637.
44. MacDonald PB, Clark P, Sutherland K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. *J Shoulder Elbow Surg* 2000;9:299-301.
45. Walch G, Boulahia A, Calderone S, Robinson AH. The 'dropping' and 'hornblower's' signs in evaluation of rotator-cuff tears. *J Bone Joint Surg Br* 1998;80:624-628.
46. Calis M, Akgun K, Birtane M, Karacan I, Calis H, Tuzun F. Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Ann Rheum Dis* 2000;59:44-47.
47. Read JW,Perko M. Shoulder ultrasound: diagnostic accuracy for impingement syndrome, rotator cuff tear, and biceps tendon pathology. *J Shoulder Elbow Surg* 1998;7:264-271.
48. Kaneko K, De Mouy EH, Brunet ME. Massive rotator cuff tears. Screening by routine radiographs. *Clin Imag* 1995;19:8-11.
49. Rahme H, Solem-Bertoft E, Westerberg CE et al. The subacromial impingement syndrome. A study of results of treatment with special emphasis on predictive factors and pain-generating mechanisms. *Scan J Rehab Med* 1998;30:253-262.
50. McFarland EG, Neira CA, Gutierrez MI, Cosgarea AJ, Magee M. Clinical significance of the arthroscopic drive-through sign in shoulder surgery. *Arthroscopy* 2001;17:38-43.
51. Adolfsson L, Lysholm J. Arthroscopy and stability testing for anterior shoulder instability. *Arthroscopy* 1989;5:315-320.
52. Nelson MC, Leather GP, Nirschl RP, Pettrone FA, Freedman MT. Evaluation of the painful shoulder. A prospective comparison of magnetic resonance imaging, computerized tomographic arthrography, ultrasonography, and operative findings. *J Bone and Joint Surg* 1991;73:707-716.
53. Westerberg CE, Solem-Bertoft E, Lundh I. The reliability of three active motor tests used in painful shoulder disorders. Presentation of a method of general applicability for the analysis of reliability in the presence of pain. *Scan Journal Rehab Med* 1996;28:63-70.
54. Nove-Josserand L, Levigne C, Noel E, Walch G. [Isolated lesions of the subscapularis muscle. Apropos of 21 cases] Les lesions isolees du sous scapulaire. A propos de 21 cas. *Rev De Chirurgie Orthopedique Et Reparatrice De l Appareil Moteur* 1994;80:595-601.

55. Gazielly DF, Gleyze P, Montagnon C, Bruyere G, Prallet B. [Functional and anatomical results after surgical treatment of ruptures of the rotator cuff. 1: Preoperative functional and anatomical evaluation of ruptures of the rotator cuff] Resultats fonctionnels et anatomiques apres traitement chirurgical des ruptures de la coiffe des rotateurs. 1ere partie: evaluation fonctionnelle et anatomique preoperatoire des ruptures de la coiffe des rotateurs. *Rev De Chirurgie Orthopedique Et Reparatrice De l Appareil Moteur* 1995; 81:8-16.
56. Lerat JL, Chotel F, Besse JL, Moyen B, Brunet Guedj E. [Dynamic anterior jerk of the shoulder. A new clinical test for shoulder instability. Preliminary study] Le ressaut dynamique anterieur de l'epaule. Un nouveau test clinique d'instabilite de l'epaule. Etude preliminaire. Rev De Chirurgie Orthopedique Et Reparatrice De l Appareil Moteur 1994;80:461-46

Chapter 6

Diagnostic evaluation of shoulder pain: A systematic review on the accuracy of signs and symptoms related to rotator cuff disorders

Abstract

Background History taking and clinical tests are commonly used to diagnose shoulder pain. It is unclear, however, whether the tests and history items used are sufficiently informative and accurate to diagnose impingement or rotator cuff tears. **Objective** To analyse the accuracy of clinical tests and history taking for instability or IAP. **Data sources** Relevant studies identified through searches of PubMed, EMBASE and CINAHL and bibliographies of known primary and review articles. **Study selection** Studies comparing the performance of history items or physical examination with a reference standard were included. Studies on fibromyalgia, fractures or systemic disorders were excluded. Of 1449 articles 35 were potentially eligible and 9 were selected. **Data extraction** Data were extracted on study population, clinical tests, reference tests and outcome. The studies' methodological quality (patient spectrum, verification, blinding, and replication) was assessed with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist. **Data synthesis** Fifteen clinical tests were found for impingement and 14 tests for rotator cuff tears. Tests showing promising results for confirming non-operable tears of the m.teres minor or m.infraspinatus were the dropping sign and Hornblower's sign, and for partial or full tears of the m.supraspinatus and/or m. infraspinatus the internal rotation lag sign (sensitivity and specificity values >0.90). For the other tests the outcomes were more heterogenous with sensitivity values ranging from 0.08 to 1.00, specificity values from 0.10 to 1.00 and the improvement of the test probability from -0.27 to 0.81 after testing. Results should be cautiously interpreted because studies were completed in selected populations in orthopedic practice, mostly assessed by the test designers and evaluated in single studies only. No accuracy studies were found for history taking or clinical tests in primary care. **Conclusion** Although history taking and clinical tests are frequently used in practice to diagnose impingement and rotator cuff tears, no evidence was found concerning the accuracy of history taking. Limited evidence is available on the accuracy of clinical tests, but only from specialised care.

1 Introduction

Shoulder pain is a common disorder in western societies and it is reported that 20 to 30% of the general population experience shoulder pain during a one-month period.[1-3] The incidence of shoulder complaints seen by the general practioner is estimated to be 1 to 5% per year[4-6] and the one-year prevalence of rotator cuff disorders in orthopaedic practice is estimated to be 2% of their total patient population.[7] The usual management of shoulder complaints in health care consists of history taking and clinical examination, followed by further diagnostic measures or treatment. It is unclear, however, whether history items and the clinical tests used are sufficiently informative and accurate to diagnose shoulder pain.
Impingement and rotator cuff tears are two well-known diagnostic categories of shoulder complaints. Both involve the rotator cuff. Impingement is seen as an initial stage of rotator cuff injuries characterized by inflammation of the tendon and rotator cuff tears might be seen as the final stage.[8-10] Several clinical tests were developed to establish these disorders, such as the tests of Neer,[8] of Hawkin,[9] of Jobe[11] and the painfull arc sign. We conducted a systematic review to assess the accuracy and usefulness of history taking and clinical tests to diagnose these disorders.

2 Methods

This review is based on the guidelines for systematic reviews of studies evaluating the accuracy of a diagnostic test.[12] MEDLINE (1966-2001), EMBASE (1980-2001) and CINAHL (1982-2001) were searched to identify relevant studies on diagnosis of shoulder pain. To retrieve all relevant publications related to diagnosing shoulder complaints in adults, the term *exp shoulder* was searched. In addition, text word searches were completed for *glenohumeral, scapula, clavicula, acromion, rotator cuff, supraspinatus, supra-spinatus, infraspinatus, infra-spinatus, serratus anterior,* and *subscapularis.* Diagnostic studies

were retrieved by exploding *sensitivity and specificity*, with additional textword searches of *specificity, false negative, screening and accuracy* based on the search strategy of Deville et al.[13] In addition, bibliographies of known primary and review articles were also examined.

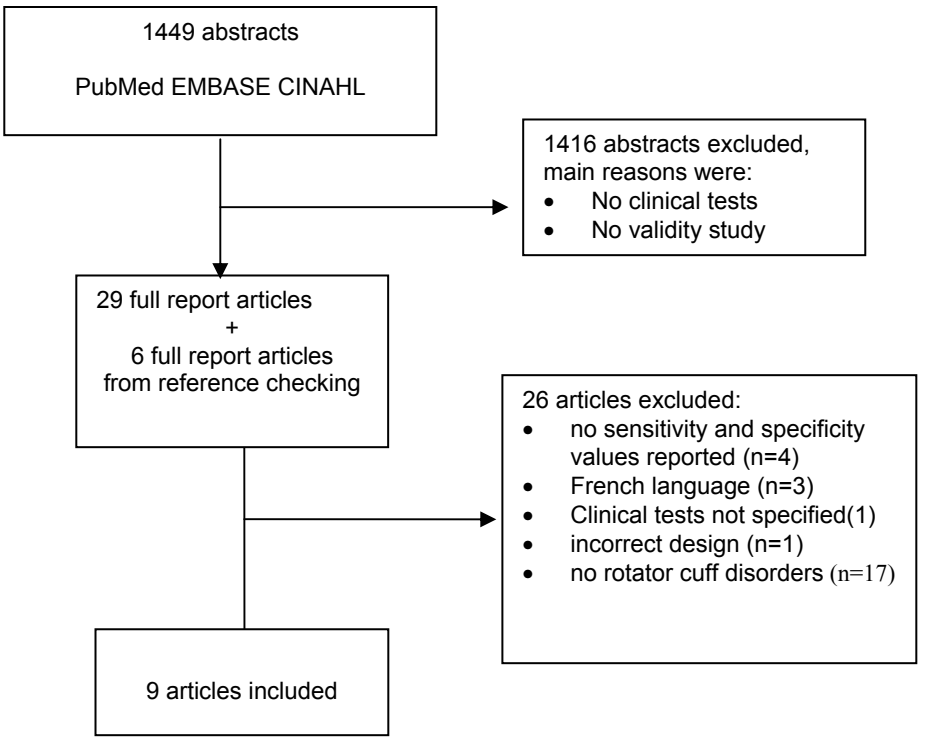
One reviewer (JJL) screened abstracts of the retrieved citations on: clinical tests, sensitivity and specificity figures, and shoulder pain. Relevant articles were obtained from the library and their reference lists were screened to find additional studies. Two reviewers (BWK and JJL) independently assessed the articles for inclusion and exclusion criteria. Minimal requirements for inclusion were full paper reports, description of clinical tests or history items for impingement or rotator cuff tears of the shoulder, a reference (gold) standard, specification of sensitivity and specificity and publication in English, Dutch or German. Studies were excluded if the diagnosis included systemic disorders such as rheumatoid arthritis and fibromyalgia, fractures or tumors. When a clinical test or symptom was assessed on its accuracy in more than one study, statistical pooling was considered. The studies should be homogenous on design, population, reference test and diagnosis, and raw data was available for the two by two table.

Methodological quality and outcome measures JJL and APV independently assessed the methods of data collection, patient selection, blinding and prevention of verification bias using the QUADAS tool, which has been developed to assess the quality of studies for diagnostic accuracy by 14 items.[14] Disagreements were resolved by consensus.

The accuracy of the diagnostic test was assessed by the outcome on sensitivity and specificity values reported in the papers. If sensitivity and specificity values were higher than 0.80, the test was seen as a good test to distinguish impingement or rotator cuff tears from other shoulder disorders. The informativeness of the diagnostic test results was determined using Bayes' theorem. Bayes' theorem considers the

positive or negative likelihood ratio of the diagnostic test in relation to the prevalence of the disease in the population under study.[15] Using this theorem one can calculate how the pre-test probability (i.e. disease prevalence) changes under the influence of diagnostic evidence into a post-test probability. The difference between post-test and pre-test probability determines the informativeness of the clinical test, and a difference of 0.3 or higher was regarded as desirable.

Figure 1 Flow chart of the selection of studies for this review



For each study, details were extracted on study population (setting, sampling, age, gender and diagnosis), clinical tests, reference tests and test performance. The primary data, on which sensitivity and specificity values were calculated, were presented when reported in the study.

3 Results

We identified 1449 studies from the electronic search, and obtained full papers for 29 of them. The reference lists of these studies revealed 6 additional studies. A total of 9 studies[16-24] met our inclusion criteria. The results of the selection procedure are shown in Figure 1.

Methodological quality The 9 studies were assessed for their methodological quality by JJL and APV. Initial disagreement occurred on 27 items (22%; kappa=0.56), but was resolved in a consensus meeting.
In 8[16-23] out of the 9 studies spectrum bias could not be excluded (Table 1). The optimal spectrum of disease, which was defined as patients visiting the orthopaedic clinic with shoulder pain, was present in one study only.[24] Blinding for the index test was unclear in 6 studies[16-18,21-23] and missing in two.[19,20] Four[16,17,21,23] studies did not describe the selection criteria for the study population and in 2 studies[18, 19] the criteria were unclear. The time between the index and reference test was unclear in 7 studies[16-19,20-22,24] and the outline of the reference test was not described in 6.[16,17,19-22]

Impingement of the shoulder Five studies evaluated 15 different clinical tests for diagnosing impingement.[17-19,23,24] Five tests were validated in more than one study. Due to the small numbers of studies per test (2 to 5) and heterogeinity among the study populations we decided not to pool the data for statistical purposes. No information was found on the accuracy of history items in diagnosing impingement. Study and patient characteristics are presented in Table 1, and test techniques are shown in the supplement at the end of this chapter.

Tabel 1 Study characteristics

Study	Inclusion criteria	Index test	Sample (n, mean age, % women)‡	Limitations¶
Prospective Design; Surgery as reference test				
Hertel et al 1996 [19]	Surgery for rotator cuff disorders; Any impairment in passive range of glenohumeral motion; Unilateral impingement	Test of Jobe External rotation lag sign Drop sign Liftoff test Internal rotation lag sign	n=100♀♂ 51 years 24%	a b d i k
Leroux et al 1995[17]	Scheduled to undergo surgery for Neer's syndroom; Chronic shoulder pain; Functional shoulder impairment	Test of Neer Test of Hawkins Test of Yocum Test of Jobe Test of Patte Lift-off test	n=55 51 years 40%	a b d l k l m
Prospective design; Arthroscopy as reference test				
Ure et al 1993[18]	Arthroscopy for shoulder pain	Test of Neer Test of Jobe Test of Hawkins 0° abduction 90° supra-spinatus test External rotation test Liftup test	n=45 42 years 14%	a b d h k
Macdonald et al 2000 [21]	Arthroscopy for shoulder pain	Test of Hawkins Test of Neer	n=85 40 years 27%	a b d f h i k m
T'jonck et al 2001 [23]	Disabling shoulder pain; Shoulder arthroscopy; <65 year Exclusion: Previous surgery of shoulder; interacting pathology	Load and shift test Apprehension test Relocation test Release test Sulcus sign Posterior stress	n=71, 45%	a b k

Study	Inclusion criteria	Index test	Sample (n, mean age, % women)‡	Limitations§
	elbow or neck	test Clunk test Test of Neer Empty can test Painful arc Resistence test external rotation Test Hawkins O'Brien's test Test of Speed Lift off test		
Prospective design Subacromial injection as reference test				
Calis et al 2000 [24]	Shoulder pain; Exclusion:Inflammatory disease; Systematic disease; Acute traumatic condition; Postoperative condition; Concom-mitant neck or elbow disorder	Test of Hawkins Test of Neer Horizontal adduction Test of Speed Test of Yergason Painful arc Drop arm	n=120 60%	c d g j m
Retrospective design; surgery as reference test				
Lyons et al 1992 [16]	Surgery for rotator cuff tears	Strength of supra-spinatus and infraspinatus and palpation of infra- and supraspinatus	N=42 40%	a b d i k m
Retrospective design; arthrography as reference test				
Litaker et al 2000 [20]	Arthrography for suspect on rotator cuff tears. Exclusion: Recent fracture of the humerus; x-ray evidence of osteo-arthritis art. Glenohumerale; Recent surgery	Supraspinatus muscular atrophy Infraspinatus muscular atrophy Elevation < 170° External rotation <70° Impingement Weakness with	n=448 57 years 37%	a f i k

Study	Inclusion criteria	Index test	Sample (n, mean age, % women)‡	Limitations⊓
		elevation Weakness with external rotation Arc of pain Expert diagnosis		
Retrospective design; surgery as reference test				
Walch et al 1998 [22]	A combined laesion of infraspinatus and supraspinatus; A full set of pre operative radiographs or CT-arthrogram; One year old tear; Full passive motion preoperatively; No rupture or subluxation of biceps tendon; No subscapularis tear; No previous surgery	Dropping sign Hornblowers sign	n=54 66 years 39%	a d i j k

‡ If data was not given for age or % females, this was indicated with '-'; ⊓ Key to limitations: a. Spectrum bias possible; b. selection criteria for waiting list not described; c. unclear if reference test is likely to correctly classify the target condition d. disease progression bias possible, time between index and reference test not described; e. partial verification bias, part of the sample did not receive the reference test; f. differential verification bias, some of the index test results were verified by a different reference standard; g. incorporation bias possible, result of the index test might be used to establish the final diagnosis; h. The execution of the index test was not described, causing problems with the study replication; i. idem h for reference test; j. unclear if clinician was blinded for the outcome of the reference test; k. idem j for the index test; m. unclear if same clinical data would be available in daily practice; n. unclear if uninterpretable or intermediate test results were reported; o. unclear if all patients entering study were accounted for (withdrawals); ∂ Patients with subscapularis tear (n=13) were excluded retrospectively

Ten clinical tests were assessed in 4 single studies using 3 different reference tests. Arthroscopy was used as a reference test in 2 studies,[18,23] surgery[17] and subacromial injection test (SIT)[24] were used in one. The results on sensitivity, specificity and information gain were heterogeneous (Table 2). High sensitivity (>0.80) and low specificity (<0.60) were reported for *the horizontal adduction test*[24] and the *O'Brien test.*[23] Low sensitivity (0.08-0.69) and high specificity (>0.80) values were found for the *drop arm test,*[24] *the resistance test in external*

rotation 0° and 90° [23] *the test of Yergason*[24] and *Jobe II.*[23] Low sensitivity and specificity values (<0.80) were found for the *test of Jobe I*[18] and *the abduction against resistance in 0° abduction.*[18] For *the test of Yocum*[17] only sensitivity (0.78) was reported. The 10 tests were not informative for the presence of impingement in the population under study, regarding the change from pre- to post-test probability (table 4). None of these tests were informative, considering the change from pre- to post test probabilities (Table 2). The probability of impingement after these tests decreased in 8 of the tests (-0.28 to –0.06) and for others there was a slight to moderate increase (0.02 to 0.28).

Table 2 Impingement; results on the outcome measurements sensitivity, pre- and post-test probability

Test Diagnosis	Sensitivity* [95% CI]	Specificity* [95% CI]	Pre P^{ρ}	Post P^{τ}
Abduction 0° [18]				
Impingement stage 2	0.69 [0.44-0.94]	0.69 [0.44-0.94]	0.29	0.39
Impingement stage 3[η]	0.78 [0.51-1.00]	0.75 [0.47-1.00]	0.20	0.38
Drop arm [24]				
Impingement	0.08 [0.03-0.13]	0.97 [0.94-1.00]	0.72	0.66
Horizontal adduction [24]				
Impingement	0.82 [0.79-0.85]	0.28 [0.20-0.36]	0.72	0.63
Test of Jobe I [18]				
Impingement stage 2	0.77 [0.54-1.00]	0.75 [0.51-0.70]	0.29	0.47
Impingement stage 3	0.55 [0.22-0.87]	0.50 [0.17-0.42]	0.20	0.18
Test of Jobe II [23]				
Impingement	0.74 (46/62) [0.64-0.84]	0.90 (9/10) [0.83-0.97]	0.86	0.86
Test of O'Brien[23]				
Impingement	0.82 (51/62) [0.73-0.91]	0.60 (6/10) [0.49-0.71]	0.86	0.64

Test Diagnosis	Sensitivity* [95% CI]	Specificity* [95% CI]	Pre P$^{\rho}$	Post P$^{\tau}$
External rotation strength test 0° [23]				
Impingement	0.69 (43/19) [0.58-0.80]	0.80 (8/10) [0.71-0.89]	0.86	0.75
External rotation strength test 90° [23]				
Impingement	0.63 (40/62) [0.52-0.74]	0.90 (9/10) [0.83-0.97]	0.86	0.84
Test of Yergason [23]				
Impingement	0.37 [0.28-0.46]	0.86 [0.80-0.92]	0.72	0.65
Test of Yocum [17]				
Impingement	0.78 [0.67-0.89]	-	1.00	-

*The numbers between brackets are the data of the 2x2 table presented in the study; $^{\rho}$ Pre P = pre test probability, is equal to the prevalence of the disease in the study population; $^{\tau}$ P post = post test probabiltiy, is equal to (LR+ x prevalence) / ((LR+ x prevalence)+1); $^{\eta}$ Impingement stage 2 is based on the classification of Neer, which impingement classified in three stage: (I) edema and hemorrhage, (II) fibrosis and tendinitis, (III) tendon degeneration bony changes and tendon ruptures.

Five clinical tests were evaluated in more than one study. Three different reference tests were used: surgery,[17,19] arthroscopy,[18,23] and subacromial injection test (SIT).[24] Study characteristics are presented in Table 1 and the results on the outcome measures in Table 3.

The test of Hawkin was assessed in 4 studies.[17,18,23,24] High sensitivity (>0.80) values were found in three studies,[17,23,24] and a moderate sensitivity (0.62) was found in one study.[18] The specificity was low to moderate (0.25-0.69) in 3 studies [18,23,24] and missing in one study.[17]

The *lift-off test*[25] has been designed to test the strength of the m.subscapularis and was evaluated in two studies.[18,19] Hertel et al.[19] used the test for diagnosing subscapularis tears and found a moderate sensitivity value of 0.62 and a specificity value of 1.00. The other study[18] assessed the value of the lift-off test for impingement not specified for the m.subscapularis and reported a sensitivity value of 0.92, and a specificity value of 0.36.[18] The test of *Neer*[8] was evaluated in four studies[17,18,23,24] for the presence of impingement. Sensitivity values ranged from 0.46 to 0.89. Specificity was computed in 3 out of

the four studies and varied from 0.31 to 0.66. *Painful arc* was assessed in two studies.[23,24] The sensitivity was 0.70 in one study,[23] and 0.33 in the other.[24] The specificity values were 0.90[23] and 0.81,[24] respectively. The test of Speed was evaluated in two studies.[23,24] The sensitivity value ranged from 0.69[24] to 0.85[23] and the specificity value from 0.56[24] to 0.80.[23]

Table 3 Impingement: results of the clinical test for impingement validated in more than one study: sensitivitity, specificity, pre- and post-test probability

Study	Diagnosis	Sensitivity* *[95% CI]*	Specificity* *[95% CI]*	Pre P^{φ}	Post P^{γ}
Test of Hawkins					
18	Impingement stage 2	0.62 *[0.36-0.88]*	0.69 *[0.44-0.94]*	0.29	0.37
18	Impingement stage 3	0.44 *[0.12-0.76]*	0.53 *[0.20-0.86]*	0.20	0.48
17	Impingement	0.87 *[0.78-0.96]*	-	1.00	-
24	Impingement	0.92 *[0.87-0.97]*	0.25 *[0.87-0.97]*	0.72	0.47
23	Impingement	0.82 *[0.73-0.91]* (51/62)	0.50 *[0.38-0.62]* (5/10)	0.86	0.58
Lift-off test					
18	Impingement stage 2	0.92 *[0.77-1.07]*	0.59 *[0.32-0.86]*	0.29	0.39
18	Impingement stage 3	0.89 *[0.69-1.09]*	0.36 *[0.05-0.67]*	0.20	-
19	Partial or complete subscapularis tear	0.62 *[0.52-0.72]* (18/29)	1.00 (24/24)	0.34	0.99
Test of Neer					
18	Impingement stage 2	0.46 *[0.19-0.73]*	0.66 *[0.40-0.92]*	0.29	0.31
18	Impingement	0.33	0.61	O.20	014

Study	Diagnosis	Sensitivity* [95% CI]	Specificity* [95% CI]	Pre P$^{\varphi}$	Post P$^{\gamma}$
	stage 3	*[0.02-0.64]*	*[0.29-0.93]*		
17	Impingement	0.89 *[0.81-0.97]*	-	1.00	-
24	Impingement	0.89 *[0.83-0.95]*	0.31 *[0.23-0.39]*	0.72	0.48
23	Impingement	0.87 *[0.79-0.95]* (54/62)	0.50 *[0.38-0.62]* (5/10)	0.86	0.60
Painful arc					
24	Impingement	0.33 *[0.25-0.41]*	0.81 *[0.74-0.88]*	0.72	0.56
23	Impingement	0.70 *[0.59-0.81]* (45/62)	0.90 *[0.83-0.97]* (9/10)	0.86	0.86
Test of Speed					
24	Impingement	0.69 *[0.61-0.77]*	0.56 *[0.47-0.65]*	0.72	0.53
23	Impingement	0.85 *[0.77-0.93]* (53/62)	0.80 *[0.71-0.89]* (8/10)	0.86	0.78

*The numbers between brackets are the data of the 2x2 table presented in the study ; $^{\varphi}$ Pre P = Pre test probability; equals the prevalence of the disease in the population under study; $^{\gamma}$ Post P = Post test probability; computed by the formula: (((Se/(1-Sp))*pre test probability)/1+ ((Se/(1-Sp))*pre test probability))

Rotator cuff tears of the shoulder No information was available on the accuracy of history items for diagnosing rotator cuff tears. We found six papers in which 20 clinical tests for the rotator cuff tears were assessed.[16,17,19-22] No test was evaluated in more than one study. Surgery was used as a reference test in 5 studies[16,17,19,21,22] and arthrography in one study.[20] Table 4 presents the results on the outcome measures of the tests.

The results on sensitivity and specificity numbers of the 20 tests were heterogeneous. Some tests scored high on both sensitivity and specificity, while others scored high on one, or low on both. Both the

dropping sign[22] and *Hornblower's sign*[22] had high sensitivity and specificity values (>0.90) for establishing non-operable tears of the m.teres minor or m.infraspinatus (fatty degeneration of the tendon>50%). The *internal rotation lag sign*[19] also scored high on sensitivity and specificity (>0.90) in diagnosing partial or full tears of the m.supraspinatus and/or m. infraspinatus. The *impingement sign*[20] and *the tests of Neer,*[20] *of Hawkin,*[21] *of Speed,*[20] of *Patte*[17] and both Jobe tests[20, 23] had high sensitivity values (>0.80) but low specificity values (0.09-0.58). Conversely, *the drop sign,*[19] *external rotation less than 70°*,[20] *the lift-off test*[20] *and external rotation lag sign*[19] scored high on specificity and low on sensitivity. Low scores on sensitivity and specificity were seen for *weakness with elevation,*[20] *weakness with external rotation,*[20] *elevation less than 170°,* [20] *infraspinatus muscular atrophy*[20] *and supraspinatus muscular atrophy.*[20] Furthermore, we found one study in which a combination of strength and pain by palpation was evaluated for the m.supraspinatus and m.infraspinatus.[16] This resulted in a high sensitivity (0.91) and moderate specificity (0.75). However, the authors did not specify which combination of symptoms and signs elicted a positive test.

The information gain for the presence of a rotator cuff tear was limited. For most of the tests the probability of a tear decreased after the test (-0.25 to –0.06) and some tests showed a slight increase (0.02 to 0.04). An exception with a change of more than 0.30 from pre- to post-test probability, were the dropping sign, Hornblower's sign, the internal rotation lag sign and the external rotation lag sign (table 6).

Table 4 Rotator cuff tears: results on the outcome measurements sensitivity, pre- and post-test probability

Test Diagnosis	Sensitivity* [95% CI]	Specificity* *[95% CI]*	Pre P[+]	Post P[∫]
Drop sign [19]				
Partial or complete infra or supraspinatus tear or both	0.21 (13/63) [0.12-0.30]	1.00 (24/24)	0.15	-[λ]

Test Diagnosis	Sensitivity* [95% CI]	Specificity* *[95% CI]*	Pre P[+]	Post P[-]
Dropping sign [22]				
Teres minor tear stage 3, or disappeared and infraspinatus tear stage 3 or 4	1.00 (10/10)	0.93 (41/44) *[0.86-1.00]*	0.19	1.00
Elevation < 170° [20]				
Partial or complete rotator cuff tear	0.30 [0.26-0.34]	0.78 *[0.74-0.82]*	0.67	0.48
External rotation <70° [20]				
Partial or complete rotator cuff tear	0.19 [0.15-0.23]	0.84 *[0.81-0.87]*	0.67	0.44
External rotation lag sign [19]				
Partial or complete infra or supraspinatus tear or both	0.70 (44/63) [0.60-0.80]	1.00 (24/24)	0.51	0.97
Hornblowers sign [22]				
Teres minor tear stage 3, 4 or disappeared and infraspinatus tear stage 3 or 4	1.00 (13/13)	1.00 (41/41)	0.24	1.00
Impingement [20]				
Partial or complete rotator cuff tear	0.97 [0.95-0.99]	0.09 *[0.06-0.12]*	0.67	0.52
Infraspinatus muscular atrophy [20]				
Partial or complete rotator cuff tear	0.36 [0.32-0.40]	0.73 *[0.69-0.77]*	0.67	0.47
Internal rotation lag sign [19]				
Partial or complete infra or supraspinatus tear or both	0.97(28/29) [0.93-1.01]	0.96 (23/24) *[0.92-1.00]*	0.53	0.93
Test of Hawkins [21]				
Rotator cuff tears	0.88 [-][K]	0.43 *[-]*[K]	0.28	0.30
Test of Jobe I [23]				
Partial or complete	0.84 (53/63)	0.58 (14/24)	0.61	0.55

Test Diagnosis	Sensitivity* [95% CI]	Specificity* *[95% CI]*	Pre P[†]	Post P[⌈]
rupture of infra or supraspinatus or both	[0.76-0.92]	*[0.48-0.68]*		
Test of Jobe II [20]				
Rotator cuff tears	0.86 [0.83-0.89]	0.50 *[0.45-0.55]*	0.62	0.51
Lift off test [20]				
Rotator cuff tear	0	0.61 *[0.56-0.66]*	0.62	-
Test of Neer [20]				
Rotator cuff tears	0.83 [][ж]	0.51 *[-]*[ж]	0.28	0.32
Test of Patte [17]				
Rotator cuff tears	0.92 [0.85-0.99]	0.30 *[0.18-0.42]*	0.62	0.45
Strength of supra-spinatus and infraspinatus and palpation of infra- and supraspinatus [16]				
Rotator cuff tears stage 1-4	0.91(31/34) [0.82-1.00]	0.75 (6/8) *[0.62-0.88]*	0.74	0.72
Supraspinatus muscular atrophy [20]				
Partial or complete rotator cuff tears	0.36 [0.32-0.40]	0.73 *[0.69-0.77]*	0.67	0.40
Test of Speed [20]				
Partial or complete rotator cuff tears	0.98 [0.97-0.99]	0.10 *[0.07-0.13]*	0.67	0.42
Weakness in elevation [20]				
Partial or complete rotator cuff tears	0.64 [0.60-0.68]	0.65 *[0.61-0.69]*	0.67	0.55
Weakness with external rotation [20]				
Partial or complete rotator cuff tears	0.76 [0.72-0.80]	0.57 *[0.52-0.62]*	0.67	0.54

[*] If data of the 2x2 table were presented in the study, the sensitivity and specificity calculations are shown in parentheses; [†] Pre P = Pre test probability; equals the prevalence of the disease in the population under study; [⌈] Post P = Post test probability; computed by the formula: (((Se/(1-Sp))*pre test probability)/1+ ((Se/(1-Sp))*pre test probability)); [λ]The post-test probability, which is based on the positive likelyhood ratio, could not be calculated due to the specificity value of 1.00; [ж] The confidence interval could not be calculated due to the indistinct numbers which have formed the sensitivity and specificity

4 Discussion

In this literature review no information was found on the accuracy of history taking. We found 29 different clinical tests, which were examined on their accuracy to establish rotator cuff disorders in specialised care. Eleven tests were evaluated for their ability to establish impingement, 14 for rotator cuff tears and 5 tests were evaluated for both disorders, although in different studies. Three tests were promising when considering their high values of sensitivity, specificity and post-test probabilities; these were the internal rotation lag sign for partial of complete rupture of infra- or supraspinatus, the Hornblower's sign and the dropping sign for non-operable tears of the teres minor and infraspinatus. We found no information on the accuracy of clinical examination of the shoulder in other health care settings such as general practice.

The outcome on accuracy of the 29 clinical tests should be interpreted with caution for four main reasons. First, the precision of the estimates of sensitivity and specificity is limited regarding their large confidence intervals as shown in table 4-6. Secondly, in most cases the informational gain of the tests on impingement and rotator cuff tears was low, even negative in a substantial number of tests, regarding the change from pre- to post-test probability. This is due on the one hand to the high prevalence rates (pre-test probability), making it difficult to add extra information on the presence of the disorder at issue, and on the other hand due to low values of sensitivity and/or specificity. Thirdly, although the methodological quality of the studies measured by the QUADAS tool was in general satisfactory, all studies scored 50% or more items positive, in most studies the validity might have been threatened. Spectrum bias may have occurred, blinding of the index test and the procedure for patient selection were unclear, and the time between the index and reference test was not mentioned. Fourthly, different reference tests (surgery, arthroscopy, arthrography, subacromial injection test) were used to evaluate the value of a clinical

test, making comparison of the results difficult. None of these reference tests are considered to be a gold standard and may therefore produce false positive and false negative results. Arthroscopy and surgery are accepted reference tests, although their sensitivity and specificity values are in detecting impingement and rotator cuff tears are unknown. This is also applies to arthrography and the subacromial injection test. The lack of reference standards, however, is common in diagnostic research and there is no perfect solution.[26] We have to deal with uncertainty in this aspect of diagnostic research.

Although we constructed a qualitative systematic review on diagnostic tests, based on recently developed methodological knowledge [12, 26] some methodological items may have influenced our results. Firstly, the cut-off points for sensitivity (>0.80) and specificity (>0.80) were arbitrarily chosen values. False positive and false negative findings are judged equally using these cut-off points. In practice one might argue that if the consequence of the test is surgery, then it is more important to have a test with a high specificity (>0.90) and a moderate sensitivity (>0.60) to prevent unnecessary surgery. Applying these cut-off values in this reviews would have resulted in one accurate test for impingement (external rotation strength test), and one additional test for rotator cuff tears (external rotation lag sign).

Secondly, we restricted our inclusion to English, Dutch and German papers. We excluded three French studies (Figure 2). However, based on their English abstract, these studies would probably also have been excluded on other criteria; 2 did not specify sensitivity and specificity values and one evaluated instability of the shoulder.

Thirdly, we had to deal with the problem of case definition, because impingement is used for different types of rotator cuff disorders. Neer classified three stages of impingement based on his theory of impingement of the structures in the subacromial space.[8] Stage 1 comprehends inflammation of the tendon with edema and hemorrhage, stage 2 covers cuff fibrosis, thickening and partial cuff tearing and in stage 3 there are full thickness tendon tears, bony change and tendon

ruptures. In some studies in this review the stage of impingement is stated[18,24] whereas in others it is used as a general term, which hampers the comparison between studies.

In conclusion: Although history taking and clinical tests are frequently used in practice, limited evidence is available concerning their accuracy. We found no information on accuracy of history taking, but identified a large number of different clinical tests aiming to establish impingement and rotator cuff tears in specialised care. The accuracy of these tests was limited, except for the dropping sign, Hornblower's sign and the internal rotation lag sign.

5 Reference List

1. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: the influence of case definition. *Ann Rheum Dis* 1997;56:308-12.
2. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M, Simmons A, Williams G. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Ann Rheum Dis* 1998;57:649-55.
3. Makela M, Heliovaara M, Sainio P, Knekt P, Impivaara O, Aromaa A. Shoulder joint impairment among Finns aged 30 years or over: prevalence, risk factors and co-morbidity. *Rheumatology (Oxford)* 1999;38:656-62.
4. Miedema HS, Rasker JJ. Het Nederlandse reumaonderzoek in de jaren tachtig en begin jaren negentig. *TNO Preventie en Gezondheid*, Leiden 1996.
5. van der Windt DA, Koes BW, de Jong BA. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis* 1995;54:959-64.
6. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands. Prevalences, consequences and risk groups, the DCM(3)-study. *Pain* 2003;102:167-78.
7. Brinker MR, O'Connor DP, Pierce P, Woods GW, Elliott MN. Utilization of orthopaedic services in a capitated population. *J Bone Joint Surg Am* 2002;84-A:1926-32.
8. Neer CS 2nd, Welsh RP. The shoulder in sports. *Orthop Clin North Am* 1977;8:583-91.
9. Hawkins RJ, Kennedy JC. Impingement syndrome in athletes. *Am J Sport Med* 1980;8:151-8.
10. Jobe FW, Moynes DR. Delineation of diagnostic criteria and a rehabilitation program for rotator cuff injuries. *Am J Sport Med* 1982;10:336-9.
11. Jobe FW, Jobe CM. Painful athletic injuries of the shoulder. *Clin Orthop* 1983:117-24.
12. Deville WL, Buntinx F, Bouter LM, et al: Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002:2;9.
13. Deville WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;53:65-9

14. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
15. Habbema JD, Eijkemans R, Krijnen P, Knottnerus JA. Analysis of data on the accuracy of diagnostic tests. In: Knottnerus JA, ed. The evidence base of clinical diagnosis. 1st ed. London: *BMJ Books*;2002:117-136.
16. Lyons AR, Tomlinson JE. Clinical diagnosis of tears of the rotator cuff. *J Bone Joint Surg Br* 1992;74:414-5.
17. Leroux JL, Thomas E, Bonnel F, Blotman F. Diagnostic value of clinical tests for shoulder impingement syndrome. *Rev Rhum Engl Ed* 1995;62:423-8.
18. Ure BM, Tiling T, Kirchner R, Rixen D. Zuverlassigkeit der klinischen Untersuchung der Schulter im Vergleich zur Arthroskopie. Eine prospektive Studie. *Unfallchirurg* 1993;96:382-6.
19. Hertel R, Ballmer FT, Lombert SM, Gerber C. Lag signs in the diagnosis of rotator cuff rupture. *J Shoulder Elbow Surg* 1996;5:307-13.
20. Litaker D, Pioro M, El Bilbeisi H, Brems J. Returning to the bedside: using the history and physical examination to identify rotator cuff tears. *J Am Ger Soc* 2000;48:1633-7.
21. MacDonald PB, Clark P, Sutherland K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. *J Shoulder Elbow Surg* 2000;9:299-301.
22. Walch G, Boulahia A, Calderone S, Robinson AH. The 'dropping' and 'hornblower's' signs in evaluation of rotator-cuff tears. *J Bone Joint Surg Br* 1998;80:624-8.
23. T'Jonck L, Staes F, De Smet L, Lysens R. The relationship between clinical shoulder tests and the findings in arthroscopic examination. *Geneeskunde En Sport* 2001;34:15-24.
24. Calis M, Akgun K, Birtane M, Karacan I, Calis H, Tuzun F. Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Ann Rheum Dis* 2000;59:44-7.
25. Gerber C, Krushell RJ. Isolated rupture of the tendon of the subscapularis muscle. Clinical features in 16 cases. *J Bone Joint Surg Br* 1991;73:389-94.
26. Knottnerus JA, Weel C. General introduction: evaluation of diagnostic procedures. In: Knottnerus JA, ed. The evidence base of clinical diagnosis. 1st ed. London: *BMJ Books*;2002:1-17

Supplement

Description of clinical tests for impingement and rotator cuff tears

Diagnostic test Shoulder position	Technique	Outcome
In the following tests the patient is sitting or standing		
Abduction 0° [18] Neutral	Clinician applies force to resist abduction	Weakness
Drop arm[24] Abducted to 90°	The patient is asked to let down the arm slowly.	Arm drops immediately with pain

Diagnostic test Shoulder position	Technique	Outcome
Drop sign[19]		
Abducted to 90° and rotated fully externally, elbow flexed in 90°.	Patient is asked to actively maintain this position as the clinician releases the wirst while supporting the elbow.	Lag or drop of minimal 5°.
Dropping sign[20]		
Externally rotated to 45° with elbow in 90° flexion	The clinician applies force to resist external rotation to the wrist	Arm drops back to the neutral position
External rotation lag sign [19]		
Elevated to 20° and near maximal external rotation.	The patient is asked to held that position	Lag or angular drop
Test of Hawkins[9]		
Forward flexed to 90° flexion	Stands facing the patient, the clinician rotates forcibly the arm internally by lowering the forearm	Pain and apprehension
Horizontal adduction[24]		
Forward flexed to 90°	Clinician forces the arm in adduction towards the other shoulder while the elbow is flexed.	Pain
Hornblowers sign 22		
Abducted to 90°, elbow flexed to 90°	Clinician applies force to resist external rotation to the wrist	Weakness in rotation
Infraspinatus muscular atrophy [20]		
Neutral	Clinician inspects scapula	Reduced muscle mass is observed if a concavity of the infraspinatus muscle is noted in conjunction with the prominence of the scapular spine
Internal rotation lag sign[19]		
Hand on the back.[λ]	The patient is asked to hold that position.	Lag or angular drop
Test of Jobe I[18]		
Forward flexed to 90° with the elbow in 90°.	The clinician pushes the hand lightly back wards or forwards	Pain[ǂ]
Test of Jobe II[23]		
Abducted to 90° and 30° horizontal adduction with the thumbs pointing downward (internal rotation)	Face the patient. The clinician pushes the patient's arms downward while asking the patient to resist pressure.	Pain
Lift off test[18]		
Hand on the back.	The clinician pulls the hand about 5 to 10 cm from the back while maintaining the 90° bend in the elbow. The patient is asked to	Hand falls back

Diagnostic test Shoulder position	Technique	Outcome
	hold the position without the examiner's help.	
Test of Neer[8]		
Neutral	The clinician fixs the scapula with one hand to prevent rotation of the scapula while passively raising the patient's arm with the other hand.[δ]	Pain
Painful arc[23]		
Abducted to 180°.	The patient is asked to actively descend the arm in the scapular plane.	Pain increases in intensity as the arm descends to 90° abduction and, is maximal between 70° and 120°.
Test of Patte[17]		
Forward flexed to 90°.	The clinician applies a force to resist external roation	Absence of pain indicates that the rotator cuff tendon is normal
External rotation strength test [23]		
Neutral position, elbow in 90° flexion.	The clinician applies force to resist external rotation	Pain or inability to resist force.
External rotation strength test in 90° abduction[23]		
Abducted to 90°, elbow flexed to 90°	The clinician applies force to resist external rotation	Pain inability to resist force.
Supra-spinatus muscular atrophy[20]		
Neutral	The clinician inspects the scapula.	Reduced muscle mass is observed superior to the supraspinatus fossa
Test of Speed[23]		
Neutral.	The clinician applies resistence to the arm while the patient is asked to elevate his arm up to 60° forward flexion.	Pain in bicipital groove area.
Strength of supra-spinatus[16]		
Abducted to 20°.	Clinician applies force to resist abduction at the wirst	Weakness
Strength of infraspinatus[16]		
Neutral. Elbow flexed to 90°.	Clinician applies force to resist internal rotation at the wrist	Weakness
Palpation of infra- and supraspinatus[16]		
Neutral	Palpate the top of the humeral head and rotated the arm and then hyperextended.	In external rotation, an interior supraspinatus tear could be felt; internal rotation revealed posterior tears and

Diagnostic test Shoulder position	Technique	Outcome
		hyperextension infraspinatus defects.[ε]
Weakness with elevation[20]		
Forward flexed to 90° and horizontal adducted to 45° with thumbs down (internal rotation).	The clinician applies downward force to the ulnar aspect of the hand.	Weakness
Test of Yergason[20]		
Neutral, elbow flexed to 90° and the forearm pronated.	The clinician applies force to resist supination at the forearm.	Pain in bicipital groove
Test of Yocum[20]		
Neutral	The patient is asked to place the hand on his or her other shoulder and to raise the elbow without elevating the shoulder.	Positive if elevation occurs
Weakness with external rotation[20]		
Elbow flexed to 90° with thumbs up, rotated internally 20°	Clinicialn applies force to resist external rotation	Weakness
In the following tests the patient is supine		
Elevation < 170° [20]		
Neutral	The clinician elevates the arme to the maximal distance.	Less than 170° elevation or the difference with the contra-lateral shoulder is more than 10°.
External rotation <70° [20]		
Neutral next to the body, elbow flexed to 90°	Using the forearm as the 'handle', the examiner rotates the humerus.	External rotation less than 70°.
Impingement[23]		
Forward flexed fully, arm against the ear.	Clinician rotates the arm internally.	Significant increase in pain

λThis is 90° elbow flexion with the shoulder in 20° elevation, 20° extension and near maximal internal rotation; ǂ if apprehension with or without pain occurs the test is positive for subluxation or dislocation; δ This reducing the space between the great tuberosity and the anterior inferior aspect of the acromion; ε It is unclear what the authors meant with 'feeling a tear'.

Chapter 7

General discussion

Summary

What is already known Shoulder complaints occur frequently in the general population. 50% to 60% of patients with shoulder complaints in primary care have persistent complaints characterised by recurrent or chronic complaints. Known risk factors for the incidence of shoulder complaints are: pushing and pulling, mental distress and obesity. The reliability of most shoulder tests to establish specific shoulder disorders ranges between poor and moderate.

What has been added Due to the great differences between diagnostic definitions, it is difficult to interpret the reported prevalence of shoulder pain in the general population. Shoulder complaints are characterised by temporal changes manifested in high rates of incidence, recurrence, and recovery. Risk factors for the onset of shoulder complaints are not necessarily risk factors for the recurrence of complaints. There is only limited evidence on the value of signs and symptoms related to specific shoulder disorders.

1 Introduction

This thesis focuses on the occurrence, course, and diagnosis of shoulder complaints. Its primary objectives are to describe the prevalence and incidence of shoulder pain in the general population, to explore the course of shoulder complaints and their determinants, and to evaluate the value of signs and symptoms related to shoulder disorders.
The purpose of this chapter is to reflect on the findings and the methodological strengths and weaknesses of the research reported here.

2 The occurrence of shoulder complaints

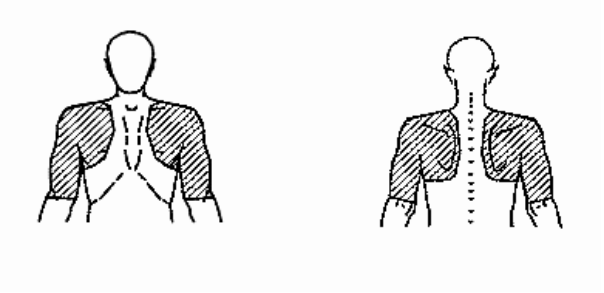
In Chapter 2, we described the range in prevalence values of shoulder complaints in the general population, and also the results of a single study on incidence. It showed that there was a strong variation in prevalence values, and that the outcome for point and life-time prevalence overlapped. Point prevalence ranged from 7 to 27%, one-month prevalence from 19 to 31%, one-year prevalence from 5 to 47% , and life-time prevalence from 7 to 67%. One-year incidence was estimated between 0.9% and 2.5% for different age categories, with the highest value for subjects aged 42 to 46 years.
The studies varied with regard to population characteristics (age and ethnicity), and applied strongly differing case definitions. Prevalence rates decreased when the case definition was restricted in terms of a minimum duration of pain or the presence of limited movements; it increased when the localisation of the pain area was enlarged. This makes it difficult to estimate the true magnitude of the prevalence of shoulder complaints.

Case definition The variety in case definitions stresses the need for consensus on the definition of shoulder complaints in research. After all,

the use of consistent case definitions will facilitate the comparability of studies, thus aiding to the generalisation of any conclusions. Different types of research questions – for example, clinical prognosis studies versus descriptive community-based studies – require presumably dissimilar definitions. If research questions are similar, standardised questionnaires can be used to synchronise the case definition. One example is the use of the Nordiq Questionnaire for musculoskeletal disorders in studies among occupational populations.[1]

When there are different types of research question, the likely solution will lie less in consensus on the case definition, but in features related to the disease. This includes consensus on the specific outline of the shoulder area, and on a report of severity and duration of the shoulder complaint. A suggestion towards consensus on the shoulder area was made by Pope et al. (1997), who studied the effect of different definitions on prevalence, and suggested using the area shown in figure 1. [2] This area covers the underlying muscles and joints that are seen as important to the movement of the shoulder.[3] Several structures of the shoulder may cause referred pain in the forearm, but as the detection of symptoms unrelated to shoulder disorders will increase if the forearm is included, we prefer to use the definition suggested by Pope and colleagues.[2]

Figure 1. Shoulder area based on the study of Pope et al.[2]



Reporting the severity and duration of shoulder complaints would give an indication of the spectrum of disease included in the study. Although the literature contains no clear definition of the severity of shoulder complaints, we feel that data, on pain, disability, and absence from work would be relevant indicators for levels of severity. Data for these three markers can be gathered using standardised instruments such as the Visual Analogue Scale for pain,[4] the 'Disability of Arm Shoulder Hand' questionnaire for disability (DASH)[5,6] and the questionnaire on absence from work by Burdorf et al.[7] Univocal and transparent communication within research and between research and practice would be greatly benefited not only if the same definition of the shoulder area were used, but also if the severity and duration of complaints were reported in different types of research, and if standardised questionnaires were used in similar types of research.[8-10]

3 The course of shoulder complaints

In Chapter 3 we explored the course of shoulder complaints over a three-year period. The results suggest that these complaints are episodic in nature rather than single events in time. In our study of workers in nursing homes and homes for the elderly, shoulder complaints were characterised by frequently recurrent episodes (> 60% per year), and a high one-year incidence (14-18%) and prevalence (32-34%). Almost one third of the people with shoulder complaints visited the general practitioner and the physiotherapist in a given year. The proportion of people who were absent from work due to shoulder complaints (for one or more episodes) varied from 14% to 24% per year.

The course of shoulder complaints was estimated within an annual timeframe, a factor that one should keep in mind when considerate the results of our study. We could not show the course of complaints within in a year, although most of our subjects had more than one episode in a single year, each of a different duration. On the basis of the one-year

incidence and recurrence, three groups could be identified: one group without complaints (46%), one with intermittent complaints (37%), and one with complaints at all three annual measurements (17%). Comparable groups were found in a smaller timeframe study on low back pain (i.e. 3-month follow-ups).[11] These similarities suggest that, despite the limitations due to our annual timeframe, our study provides valuable information on the nature of shoulder complaints over time.

Episodic nature Given the high frequency of recurrence, shoulder complaints seem to be episodic in nature rather than characterised by acute self-limiting or persistent complaints. More specific information on this episodic nature is therefore called for. Firstly, greater insight into the course of an episode would be valuable, especially with regard to the variation in pain, the restrictions in motion, and the disability perceived from day to day. Secondly, it would be useful to know how frequently episodes occur in a certain timeframe (e.g. within one-year), and whether they differ in localisation, duration, nature and severity. Thirdly, it is important to evaluate whether these patterns are sustained over a number of years. Finally, within the episodic nature of shoulder complaints, greater understanding is needed of the factors that cause the complaints to flare up, of the factors leading sufferers to seek health care, and of the factors that lead to a period of sick leave. Such information would facilitate a better choice, timing, and duration of interventions, and would also indicate whether a wait-and-see policy is enough.

Risk factors In Chapter 4 we studied the influence of personal and work-related risk factors on the incidence and recurrence of shoulder complaints, and asked whether the risk factors for incidence were similar to those for recurrence. The statistical analysis suggested that, to some extent, these risk factors are dissimilar. Obesity and some work-related factors were associated with the incidence of shoulder complaints, but not with their recurrence. Only ‘manual material-

handling' was associated with both incidence and recurrence. One other study has evaluated the risk factors for both incidence and recurrence of shoulder complaints: it found that incidence was associated with age, mental distress, and physically strenuous work, while recurrence was related only to age and overload at work.[12] These results suggest that risk factors for incidence cannot be considered as apparent risk factors for recurrence. This raises the question of why risk factors for incidence and recurrence may be different. In our opinion, it is more likely that specific risk factors causing the development of a shoulder complaint will also play a role in initiating a recurrent episode of this complaint. Choices in the study design may therefore have influenced the differences we observed. Firstly, reference groups differed because incident cases in this type of research are compared with healthy subjects, and because recurrent cases are compared to recovered cases. Therefore, self-reported exposures may have depended on the health status of the subjects, and may thus have influenced the outcome of the study. However, in our study, cases and healthy subjects did not differ in their estimation of physical and psychosocial work exposures, suggesting that the reported exposures were not biased by their health status.

Secondly, it is also possible that, due to our annual timeframe, we did not pick up the variation in certain risk factors. Recurrent episodes may have been caused by temporarily high levels of physical or psychological stress in work or personal life. As noted by Mittleman et al, it is not easy to measure these variations, which require accurate measurements of exposures and covariates during the etiologically relevant time periods in the follow-up period of a cohort study.[13] Because a normal cohort study design does not make it easy to answer such a time-related question, it would be useful to conduct a case-crossover analysis within a cohort study. Analysing events with a case-crossover design requires well-defined and easy-to-determine events. Prospectively, data should be gathered on stable and variable risk factors and complaints. Risk patterns for new episodes would be

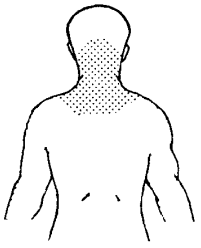
analysed the usual way, while information on recurrent episodes would be collected on the transient risk factors immediately preceding the episode. In this approach, each individual would form his or her own stratum.

Definition of Episodes Episodes of recurrent shoulder complaints are not well defined in literature. In a recent review of low back pain it was suggested that an episode of low back pain may be determined when a complaint persisted for at least 24 hours, preceding and following a month free of complaints.[14] This definition might work as well for shoulder complaints, but limits its use in a case-crossover analysis. Study subjects have to be aware when the complaints have occurred, whether these were present for at least 24 hours, and when the complaint ended. Given this definition, monthly measurements are probably necessary to accurately determine the risk factors related to recurrent events.

Neck complaints The course of neck complaints was also explored in this longitudinal study. This corresponded strongly with the course of shoulder complaints, which was not surprising, as 50% to 60% of the cases with shoulder complaints also had neck complaints. Unfortunately, due to our annual study frame, we do not know whether these episodes of neck and shoulder complaints occurred together or separately. However, the concurrence of shoulder and neck complaints raises the question of whether it is justified to make a distinction between them. In our research, this question is difficult to answer. In previous research some authors have chosen to combine neck and shoulder into one complaint they identify as neck-shoulder complaints.[15-19] Others preferred to assess neck and shoulder complaints separately. However, drawings of the two areas sometimes contain an overlap, and, when questioned, subjects may not be sufficiently able to differentiate between shoulder and neck complaints.

When studying risk factors, it is therefore uncertain whether one should combine or separate complaints. Thought the eventual choice depends strongly on the research question, we feel it is wise to separate neck and shoulder complaints in accordance with clinical practice, because their treatment differs. In population-based research this can be achieved by using the definition of the shoulder area as suggested in Figure 1, and the definition of the neck area as suggested in the Nordiq Questionnaire[1] (see Figure 2). In patient-based studies it may be possible to separate neck complaints from shoulder complaints by means of clinical examination. However, it is likely that some subjects will have both complaints. So, it would be advisable to report the occurrence of both complaints in studies on shoulder or neck complaints.

Figure 2 Neck area based on the Nordiq Questionnaire[1]



4 The diagnosis of shoulder disorders

There is only limited evidence on the value of signs and symptoms related to specific shoulder disorders. In Chapters 5 and 6 we described 26 studies on the sensitivity and specificity of 50 clinical tests. Sixteen of these assessed signs related to instability or intra-articular pathology (e.g. labral tears), 8 evaluated signs of impingement or rotator cuff tears, and one evaluated signs of both instability and rotator cuff tears. Eleven tests were valid in study populations consisting of patients who were on the list for shoulder surgery or arthroscopy. The relocation and

anterior release test added information on the presence of instability, and the Biceps load I and II test, the Internal Rotation Strength test, and the test of Mimori were informative for diagnosing intra-articular problems. None of the tests on impingement seemed to be valid for the presence of impingement. The Internal Rotation Lag Sign was valid for rotator cuff tears, and the Hornblower's sign and the Dropping sign were valuable for non-operative rotator cuff tears (fatty degeneration of the tendon >50%). The validity of history items related to specific shoulder disorders remains unclear.

While the term 'diagnosis' implies that certain clinical presentations lead to the recognition of a specific disorder, it is doubtful whether this is the case with shoulder complaints. We found limited evidence that specific shoulder disorders can accurately be established on the basis of individual signs. While two studies tried to cluster sign and symptoms related to shoulder complaints, they did not reveal specific discriminatory groups of shoulder pain.[20,21]

There are also questions on the reliability of shoulder tests. In the discrimination of shoulder pain by two examiners, in both primary and specialised care, the inter-observer agreement was only poor to moderate.[22-24]

Diagnostic classification The diagnostic classification of shoulder complaints is a recurrent point of discussion. This is reflected in the nomenclature, which is based on different etiological, pathophysiological, and clinical classifications and often used side by side.[10]

One of the classification systems is based on the underlying aetiology, such as impingement. This implies a clear understanding of the pathophysiological mechanisms. However, even though we know that impingement is due mostly to compression of structures in the subacromial space, it is unclear which mechanisms cause this compression. Several potential external factors are cited, such as morphology of the coracoacromial arch, tensile overload, repetitive use,

and kinematic abnormalities, but so, too, are intrinsic factors, such as altered tendon vascular supply, microstructural collagen fibre abnormalities, and regional variation in material property.[25,26] However, there is little quantitative information regarding the relative significance of these individual factors.

A second classification system uses the pathophysiological expressions in body tissue, such as capsulitis, tendinosis, bursitis, rotator cuff tears and labral tears. These expressions are based on histological findings at surgery, on arthroscopy and in laboratory cadaver studies, in which the pathophysiological expressions were retrospectively connected to the clinical presentations in patients operated upon for shoulder pain.[25,26] A recent example of this is the clinical presentation of labral tears. Renewed interest in these were caused by developments in arthroscopy in the shoulder, which brought the possibility to repair the labral tears. Snijder et al.[27,28] related the labral tear to pain deep in the shoulder; this was accompanied by clicking or popping sounds. Others have argued against this, because there are several other disorders which may have similar clinical presentation, such as pathologic conditions of the rotator cuff.[29,30]

The third classification system is based on abnormalities in the range of motion, such as instability or frozen shoulder. This seems a logical classification, as one can easily determine the range of motion. However, the major problem is the definition of these disorders.[8,9] When a shoulder dislocates and the patient presents with disabling pain and his arm fixed to his body, the diagnosis is easy to establish. More often, however, subtle forms of instability may play a role in the shoulder complaints, though these are far less easy to establish. The same applies to frozen shoulders. This diagnosis is easy to establish when the range of shoulder motion is near to zero. However, more often the range of movement is not fully restricted, which makes it difficult to establish the diagnosis.

The three classification systems are not mutually exclusive. The term impingement, of the first classification system, encloses disorders of the

bursa and rotator cuff tendons, which belong to the second classification. And, the term frozen shoulder of the third classification system is also known as capsulitis adhesive, which belongs to the second system. This causes difficulties in univocal and transparent communication about specific shoulder disorders.

Causes, manifestation, prognostic profiles Despite the lack of a clear and valid classification system, it is likely that patients with shoulder complaints comprise several subgroups with different causes, manifestations, and prognostic profiles. It is known from histological research in cadavers and animals that damage can occur in all structures of the shoulder (e.g. tendon, bone, capsule).[25] These structures (except the labrum) are innervated with nociceptive fibres,[31] which means that they can all be a potential source of shoulder pain. It is, however, difficult to relate shoulder symptoms and signs to histological characteristics (such as tendinosis and capsulitis) in living human beings.

Because surgery, biopsy, or arthroscopy will not be indicated in most cases of shoulder pain, there is no way of verifying the presence of tissue damage. For these patients it would be more appropriate to use a simpler model based on features which can be established in practice, such as the localisation, duration and severity of pain, restrictions in movement and strength, level of disability and prognostic profiles. These features can be used in the treatment of shoulder complaints or in the evaluation of the treatment. An exception might be any shoulder disorder that indicates surgery or arthroscopy, such as rotator cuff tears and labral tears; for these disorders it might be relevant to search for the specific combination of signs and symptoms related to the tissue damage.

In some cases shoulder complaints are caused by other underlying pathologies such as neurological or vascular disorders, neoplasms, and referred pain from internal organs. Although it is not clear in which way these disorders present – and they may be easy to confuse with other

disorders of the shoulder – there are several indications in terms of red flags: constant progressive non-mechanical pain, history of drug abuse, cancer or HIV, weight loss, violent trauma, and widespread neurological signs and symptoms. [32] A less severe source of shoulder complaints is referred pain from the neck. In cases with referred pain from the neck, the shoulder pain cannot be provoked by movements of the shoulder, but is easily provoked by movements of the neck.[33]

Methodology The limited evidence found for the validity of clinical tests may partly be explained by the methodology used in the primary studies. Most studies in our systematic review have used standardised methods used in diagnostic research. Regarding the outcome on the QUADAS (quality assessment for diagnostic accuracy studies),[34] the studies were of reasonable quality, though they were limited by the highly selected populations, and the gaps in descriptions of the selection criteria, reference test, and blinding procedures.
The use of highly selected populations in the primary studies limits the extent to which their outcomes can be generalised. These are valuable only for these highly selected patient populations. Moreover, the use of such populations might also have introduced difficulties in diagnosing shoulder disorders. Most of the people undergoing surgery or arthroscopy have had complaints for over 12 months, which, according to the definition of the International Association for the Study of Pain (IASP), means that these complaints are chronic. In their study on diagnosing shoulder complaints (mean duration of complaint 25 months), Norregaard et al.[9] suggested that diagnostic studies in patients with a shorter history may lead to more valid and reliable results of history and clinical examination. In chronic complaints it is more difficult to find the primary source of pain, and a significant role may be played by other factors. In the research on chronic low back pain, associations have been found with fear avoidance, self-efficacy beliefs, and coping strategies.[35-37] These factors may also be relevant to longstanding shoulder complaints. It is therefore likely that patient

populations with longstanding complaints are less valuable in assessing the value of the history and clinical examinations.

Although appropriate study designs were used to establish the validity of individual signs, the clinical presentation of shoulder complaints will seldom be investigated using an individual finding. In general, a combination of signs and symptoms is used to establish a specific shoulder disorder. Recently, Murell and Walton compared the value of 23 tests for the presence of rotator cuff tears.[38] They found that the combination of supraspinatus weakness, weakness in external rotation, and a positive impingement sign raised the probability of a rotator Cuff tear to 98%. This type of research meets the reality of clinical practise (i.e. combining the findings of different signs and symptoms). It is therefore advisable to extend this approach to specific disorders, such as partial rotator cuff tears and labral tears, which can be verified by surgery or arthroscopy.

Fortunately, knowledge of diagnostic research methodology has increased over the last few years. Compared to effectiveness studies, however, it is still in its infancy. Although previous diagnostic studies used sensitivity and specificity figures to express the value of a test, these outcome measures can not be used on their own: it is their combination that matters.[39] Three diagnostic outcome measure combine sensitivity and specificity: the likelihood ratio, the post-test probability based on Bayes theorem, and the diagnostic odds ratio. The latter summarises the diagnostic information, but does not provide the specific values of sensitivity, specificity and likelihood ratios. We therefore preferred to use the post-test probabilities based on Bayes theorem, because this combines the knowledge of the likelihood ratio with the pre-test probability to a post-test probability.[39] The magnitude of change from pre-test to post-test probability reflects the informativeness of the diagnostic test result.[39] These measure is, however, not commonly used, and its meaning is difficult to general readers. In Chapter 5 we therefore described only the likelihood ratios.

5 Conclusion

Accurate identification of cases with shoulder disorders is a recurrent point in this thesis. In the systematic review on the incidence and prevalence of shoulder pain, we have seen that differences between the case definitions of shoulder pain hampered our overall conclusions. It also played a role in the description of the course of shoulder complaints. In the absence of an accurate system for identifying cases with a specific shoulder disorder, the management of these disorders will not be optimal.

6 Recommendations for research

In this thesis we have discussed several gaps in the knowledge of shoulder complaints related to their epidemiology and diagnosis. Some important recommendations for future studies in the epidemiology and diagnosis of such complaints can be derived from the previous chapters. In order of appearance, they are:

1. To contribute to the generalisability of research on shoulder complaints in different types of research, we recommend the use of the definition of the shoulder area proposed by Pope et al.[2] and measuring the levels of severity and duration of the complaint by standardised questionnaire on pain, disability and absence from work. In similar types of research standardised questionnaires can be used to synchronise the case definition.
2. The effects and sustainability of therapeutic interventions in shoulder complaints depend on the natural course of shoulder complaints. Further research on the course of shoulder complaints should therefore take account of the episodic nature of these complaints.
3. To prevent recurrent episodes of shoulder complaints, future studies should distinguish between risk factors for their onset, and risk factors for their recurrence.

4. To gain greater insight into the interrelationship between neck and shoulder complaints, the overlap between these complaints should be investigated.
5. For shoulder disorders that cannot be verified with reference methods such as arthroscopy or surgery, a classification system based on anatomical or physiological descriptions should not be used. A more pragmatic classification could be based on the localisation, duration and severity of pain, restrictions in movement and strength, level of disability and prognostic profiles. The usefulness of such a classification system should be investigated in relation to patient outcomes.
6. For shoulder disorders that can be verified with reference methods such as arthroscopy or surgery, a classification based on histological findings may be appropriate, provided their symptoms and clinical signs are sufficiently different from other shoulder disorders.
7. The use of history items to determine specific shoulder disorders is common in medical practise. However, because no scientific evidence is available, research is required to assess the value of history items.

7 Consequences for clinical practice

In clinical practice shoulder disorders are often approached from anatomical, physiological, and pathological viewpoints. There is limited evidence on identifying different disorders on the basis of clinical presentations of pathophysiological processes. While in some cases the underlying pathology can be verified in surgery, there is no valid clinical presentation in terms of signs and symptoms.
As suggested by other researchers[22-24] it would be appropriate to approach shoulder disorders in terms of their presentation (e.g. pain, range of motion, disability). This would fit into current treatment for the shoulder, which is based mainly on reducing pain and improving

function in daily life (e.g. in work and sports). Due to the different treatments provided by the general practitioner, physiotherapist and orthopaedic surgeon, it is likely that the categorisations per setting will differ.

The clinical guideline on shoulder complaints of the Dutch College of General Practice makes a distinction between shoulder patients with restrictions in motion and those without.[40] For the physiotherapist this is presumably a good first step, but more refinement is needed in guidance on the use of exercises and appropriate mobilisation techniques. In orthopaedic practice categorisation should guide the use of diagnostic modalities and therapeutic opportunities, such as shoulder-imaging, injection, surgery or arthroscopy.

8 Reference List

1. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sorensen F, Andersson G, Jorgensen K. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 1987;18:233-237.
2. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: the influence of case definition. Ann Rheum Dis 1997;56:308-312.
3. van der Helm FC. A finite element musculoskeletal model of the shoulder mechanism. J Biomech 1994;27:551-569.
4. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983;17:45-56.
5. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome meaure: the DASH (disabilities of the arm, shoulder and hand). The Upper Extremity Collaborative Group (UEGG). Am J Ind Med 1996;29:602-608.
6. Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionaires. A systematic review. Ann Rheum Dis In press.
7. Burdorf A, Post W, Bruggeling T. Reliability of a questionnaire on sickness absence with specific attention to absence due to back pain and respiratory complaints. Occup Environ Med 1996;53:58-62.
8. Van Eerd D, Beaton D, Cole D, Lucas J, Hogg-Johnson S, Bombardier C. Classification systems for upper-limb musculoskeletal disorders in workers: a review of the literature. J Clin Epidemiol 2003;56:925-936.
9. Norregaard J, Jacobsen S, Kristensen JH. A narrative review on classification of pain conditions of the upper extremities. Scand J Rehabil Med. 1999;31:153-164.

10. Katz JN, Stock SR, Evanoff BA, Rempel D, Moore JS, Franzblau A, Gray RH. Classification criteria and severity assessment in work-associated upper extremity disorders: methods matter. Am J Ind Med 2000;38:369-372.
11. Smedley J, Inskip H, Cooper C, Coggon D. Natural history of low back pain. A longitudinal study in nurses. Spine 1998;23:2422-2426.
12. Miranda H, Viikari-Juntura E, Martikainen R, Takala EP, Riihimaki H. A prospective study of work related factors and physical exercise as predictors of shoulder pain. Occup Environ Med 2001;58:528-534.
13. Mittleman MA, Maldonado G, Gerberich SG, Smith GS, Sorock GS. Alternative approaches to analytical designs in occupational injury epidemiology. Am J Ind Med 1997;32:129-141.
14. De Vet HC, Heymans MW, Dunn KM, Pope DP, van der Beek AJ, Macfarlane GJ, et al. Episodes of low back pain: a proposal for uniform definitions to be used in research. Spine 2002;27:2409-2416.
15. Smedley J, Inskip H, Trevelyan F, Buckle P, Cooper C, Coggon D. Risk factors for incident neck and shoulder pain in hospital nurses. Occup Environ Med 2003;60:864-869.
16. Bertilson BC, Grunnesjo M, Strender LE. Reliability of clinical tests in the assessment of patients with neck/shoulder problems-impact of history. Spine 2003;28:2222-2231.
17. Andersen JH, Kaergaard A, Mikkelsen S, Jensen UF, Frost P, Bonde JP, Fallentin N, Thomsen JF. Risk factors in the onset of neck/shoulder pain in a prospective study of workers in industrial and service companies. Occup Environ Med 2003;60:649-654.
18. Josephson M, Ahlberg G, Harenstam A, Svensson H, Theorell T, Wiktorin C,Vingard E. Paid and unpaid work, and its relation to low back and neck/shoulder disorders among women. Women Health 2003;37:17-30.
19. Feveile H, Jensen C, Burr H. Risk factors for neck-shoulder and wrist-hand symptoms in a 5-year follow-up study of 3,990 employees in Denmark. Int Arch Occup Environ Health 2002;75:243-251.
20. Winters JC, Groenier KH, Sobel JS, Arendzen HH, Meyboom-de Jongh B. Classification of shoulder complaints in general practice by means of cluster analysis. Arch Phys Med Rehabil 1997;78:1369-1374.
21. Groenier KH, Winters JC, de Jong BM. Classification of shoulder complaints in general practice by means of nonmetric multidimensional scaling. Arch Phys Med Rehabil 2003;84:812-817.
22. de Winter AF, Jans MP, Scholten RJ, Deville W, van Schaardenburg D, Bouter LM. Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. Ann Rheum Dis 1999;58:272-277.
23. van der Windt DA, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. Ann Rheum Dis 1995;54:959-964.
24. Norregaard J, Krogsgaard MR, Lorenzen T, Jensen EM. Diagnosing patients with longstanding shoulder joint pain. Ann Rheum Dis 2002;61:646-649.
25. Soslowsky LJ, Carpenter JE, Bucchieri JS, Flatow EL. Biomechanics of the Rotator Cuff. Orthop Clin N Am 1997;28:17-29.
26. Soslowsky LJ, Thomopoulos S, Tun S, Flanagan CL, Keefer CC, Mastaw J, Carpenter JE. Neer Award 1999. Overuse activity injures the supraspinatus tendon in an animal model: a histologic and biomechanical study. J Shoulder Elbow Surg 2000;9:79-84.

27. Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ.SLAP lesions of the shoulder. Arthroscopy 1990;6:274-279.
28. Snyder SJ. Labral lesions (Non-Instability) and SLAP lesions. In: Snyder SJ. Shoulder arthroscopy. 1994, New York; McGraw-Hill:115-131.
29. McFarland EG, Kim TK, Savino RM. Clinical assessment of three common tests for superior labral anterior-posterior lesions. Am J Sports Med 2002;30:810-815.
30. Stetson WB, Templin K. The Crank test, the O'Brien test, and routine magnetic resonance imaging scans in the diagnosis of labral tears. Am J Sport Med 2002;30:806-809.
31. Vangsness CT, Ennis M, Taylor JG, Atkinson R. Neural anatomy of the glenohumeral ligaments, labrum and subacromial bursa. Arthroscopy 1995:180-184.
32. Brox JI. Regional musculoskeletal conditions: shoulder pain. Best Pract Res Clin Rheumatol 2003;17:33-56.
33. Bogduk N. The anatomy and pathophysiology of neck pain. Phys Med Rehabil Clin N Am 2003;14:455-472.
34. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.
35. Vlaeyen JW, Linton SJ. Fear-avoidance an its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317-332.
36. Kuch K. Psychological factors and the development of chronic pain. Clin J Pain 2001;17:S33-38.
37. Linton SJ. A review of psychological risk factors in back and neck pain. Spine 2000;25:1148-1156.
38. Murell GA, Walton JR. Diagnosis of rotator cuff tears. The Lancet 2001;357:769-70.
39. Habbema JD, Eijkemans R, Krijnen P, Knottnerus JA. Analysis of data on the accuracy of diagnostic tests. In: The evidence base of clinical diagnosis. First edition. Ed. Knottnerus JA. 2002 London;BMJ Books: 117-136.
40. NHG standard shoulder complaints: http://nhg.artsennet.nl/upload/ 104/guidelines2 /E08.htm.

Summary

Summary Pain and restricted motion are prominent symptoms in shoulder complaints, which can impede the use of the arm in daily life, for example when combing the hair or performing personal hygiene, or any other task requiring free movement of the shoulder. Although shoulder complaints are a common musculoskeletal complaint, the estimated prevalences of it vary considerably. Little is known about the course of shoulder complaints and the risk factors related to them. In health care, the management of such complaints is based on history and clinical examination, whereby shoulder problems are classified in terms of more specific disorders, such as instability, arthritis or rotator cuff disorders. However, the validity of signs or symptoms related to these specific disorders is unclear.

Chapter 2 presents a systematic review of prevalence and incidence studies in the general population. The prevalence of shoulder complaints was estimated in 13 studies, with different estimates ranging from 5% to 67% with overlap between point and life-time prevalence. The incidence of shoulder complaints found in a single study ranged from 0.9% to 2.5%, depending on age category, with the highest estimate for the 42-46 year age group. The prevalence of shoulder-arm pain was estimated in five studies. The one-month and life-time prevalence was estimated at around 30% and the presence of chronic shoulder-arm complaints was estimated to be 8% and 20%. The outcome of the prevalence was strongly determined by the case definition used. Prevalence decreased when the case definition was restricted in terms of duration of the complaints or required also restrictions in movement, and increased when the localisation of the area of pain was enlarged. The results also suggested that prevalence increased with age and that it was higher in women.

Chapter 3 describes a cohort study on the course of shoulder and neck complaints over a three-year period, in which we estimated the 12-month prevalence, incidence and recurrence among workers of nursing homes and homes for the elderly. Over three years, 346 subjects completed a questionnaire each year. The 12-month incidence rates for neck and shoulder complaints were 16%-18%; 12-month prevalence rates were roughly twice as high, and 12-month recurrence rates were

approximately twice the prevalence rates (60-65%). Each year, 21%-38% of the subjects sought medical care for neck or shoulder pain. Between 13%-24% of the subjects reported periods when this pain made them absent from work each year. Our results suggest that, in most subjects, neck and shoulder complaints run a recurrent course characterised by a strong variation in occurrence, rather than an acute, self-limiting course.

In Chapter 4 we compare the risk factors for incidence and recurrence of shoulder and neck complaints, and study them for similarities and differences. We studied the influence of age, gender, obesity (Body Mass Index >30 kg/m^2), physical and psychosocial workload, general health and need for recovery in the cohort described in Chapter 3. Obesity was related to the incidence of shoulder complaints in the multivariate model, adjusted for age and gender. The incidence of neck complaints was higher in obese people, in those who worked in awkward postures, and those in poor/fair general health. Recurrence of both neck and shoulder complaints was associated with chronic complaints at baseline, but not with work-related risk factors. These results suggest that there are differences between risk factors for incidence and recurrence for both neck and shoulder complaints.

Chapters 5 and 6 review the literature on the validity of history and clinical examination for diagnosing specific shoulder disorders. Our literature search retrieved 25 studies containing data on the validity of 50 clinical tests for four disorders: instability, labral tears, impingement and rotator cuff tears. No data was available on other disorders or on the validity of history items. Regarding the outcome on the QUADAS (quality assessment for diagnostic accuracy studies),[34] the studies were of reasonable quality, though they were limited by the highly selected populations, and the gaps in descriptions of the selection criteria, reference test, and blinding procedures.

Chapter 5 presents the results of the clinical tests on instability in five studies, and the results of thirteen studies evaluating labral tears or other intra-articular pathology. The different tests were compared on the basis of the likelihood ratios. To establish instability, the relocation test and the anterior release test favour the diagnosis. Five manoeuvres

were less useful: the apprehension test, clunk test, release test, load and shift tests, and sulcus sign. The most promising tests for establishing labral lesions were the biceps load I and biceps load II tests, the pain-provocation test of Mimori, and the internal rotation resistance strength test.

Chapter 6 reviews the results of the validity studies on the clinical tests for impingement and rotator cuff tears. The different tests were compared for outcome with regard to sensitivity, specificity, and post-test probability based on Bayes theorem. To be valuable as a single diagnostic instrument, the test should have arbitrary sensitivity and specificity over 0.80; post-test probability should be 0.30 higher than pre-test probability. None of the 15 clinical tests on impingement fulfilled these three criteria, suggesting that they are not valuable as single diagnostic tests. Three out of 19 clinical tests for rotator cuff tears seemed to be adequate as single diagnostic tests. The results of Hornblower's sign and the dropping sign suggested that both tests can establish non-operable tears of the m.teres minor and m.infraspinatus. The internal rotation lag sign indicates the presence of partial or complete tears of the m.infraspinatus or m.supraspinatus.

Chapter 7 reflects on the findings of this thesis, in which accurate identification of cases with shoulder disorders is a recurrent point. Differences in case definition of shoulder complaints hampered the overall conclusion in the systematic review on the incidence and prevalence of shoulder complaints in the general population. It also played a role in the determination of the course of shoulder complaints. Because there is no accurate system for identifying cases with a specific shoulder disorder, the management of these disorders will not be optimal. These points of discussion stressed the need for consensus on the case definition of shoulder complaints, the definition of episodes to describe the course of complaints and the need for a classification system of shoulder disorders which meet the reality of today's clinical practice. Some suggestions on meeting these needs are stated.

Samenvatting

Samenvatting Pijn en bewegingsbeperking zijn kenmerkende symptomen bij schouderklachten. Ze kunnen hinder veroorzaken bij het uitvoeren van dagelijkse activiteiten zoals het kammen van het haar, persoonlijke hygiëne en tal van andere activiteiten die volledige bewegingsvrijheid van de schouder vragen. Hoewel bekend is dat schouderklachten vaak voorkomen, is de precieze frequentie van deze klachten in de open populatie onduidelijk. Daarnaast is er weinig kennis over het beloop van schouderklachten en daarmee samenhangende risicofactoren. Schouderklachten worden in de praktijk ingedeeld in specifieker afwijkingen, zoals instabiliteit, artritis of rotator cuff aandoeningen, aan de hand van anamnese en lichamelijke onderzoek. Over de validiteit van de anamnese en lichamelijk onderzoek om deze specifieke aandoeningen te diagnosticeren is weinig bekend.
In hoofdstuk 2 wordt de frequentie van schouderklachten in de open bevolking beschreven aan de hand van de uitkomst van een systematische review. Dertien studies beschreven de prevalentie van schouderklachten. De prevalentie varieerde tussen 5% en 67% voor de verschillende perioden, waarbij de punt-prevalentie en lifetime-prevalentie elkaar overlapten. Eén studie bevatte gegevens over de incidentie van schouderklachten. Deze werd geschat op 0,9% tot 2,5% voor verschillende leeftijdscategorieën. De hoogste incidentie werd gemeten in de leeftijdsgroep 42-46 jaar. De prevalentie van schouder-bovenarm-klachten (lifetime en 1-maands) werd geschat rond de 30% in drie studies. De prevalentie van chronische schouder-bovenarm-klachten werd geschat tussen de 8% en 20% in twee studies. De overlap in prevalentie cijfers werd sterk bepaald door de definitie van de klacht. Wanneer de definitie naast de aanwezigheid van klachten ook eisen stelde aan de duur of de aanwezigheid van bewegings-beperkingen nam de prevalentie af. De prevalentie nam toe wanneer het gebied waarin de klachten zich konden voordoen werd uitgebreid. De resultaten van de studies suggereerden tevens een hogere prevalentie voor vrouwen en een stijging naarmate de leeftijd van de onderzoekspersonen toenam.

Hoofdstuk 3 bevat gegevens over de 12-maandse incidentie, prevalentie en recidivering van schouder- en nekklachten in cohort studie onder werknemers uit verzorgings- en verpleeghuizen. Over een periode van drie jaar vulden 346 werknemers jaarlijks een vragenlijst in. De 12-maands incidentie van nek- en schouderklachten werd geschat op 16-18%, de 12-maands prevalentie op 32-36%, en het recidief percentage op 60-65%. Medische hulp werd jaarlijks gezocht door 21-38% van de mensen met klachten, en 13% tot 24% van mensen verzuimden vanwege nek- of schouderklachten. De bevindingen van deze studie suggereren dat schouder- en nekklachten een sterk recidiverend beloop kennen, in tegenstelling tot de in de literatuur vaak genoemde acute kortdurende karakter van deze klachten.
In hoofdstuk 4 worden de risicofactoren voor de 12-maandse incidentie vergeleken met de factoren voor recidivering van schouder- en nekklachten. Aan de hand van het cohort zoals beschreven in hoofdstuk 3 werd de invloed bestudeerd op het ontstaan en recidivering door leeftijd, geslacht, zwaarlijvigheid (Body Mass Index >30 kg/m2), fysiek belastende factoren in het werk, psychosociale belastende factoren in het werk, algemene gezondheid, herstelbehoefte en sport. In de multivariate analyse, gecorrigeerd voor leeftijd en geslacht, bleek de incidentie van schouderklachten gerelateerd aan zwaarlijvigheid. De incidentie van nekklachten was geassocieerd met zwaarlijvigheid, werken in belastende houdingen en een slechte tot matige algemene gezondheid. Het recidiveren van zowel schouder- als nekklachten was gerelateerd aan langdurige klachten (> 3 maanden) in het jaar ervoor. De resultaten uit deze studie suggereren dat er verschillen zijn tussen de risicofactoren voor de incidentie en de recidivering van schouder- en nekklachten.
Hoofdstuk 5 en hoofdstuk 6 bevatten de resultaten van de systematische review naar de validiteit van anamnese en lichamelijk onderzoek bij de diagnostiek van specifieke schouderklachten. De systematische literatuurstudie leverde 26 studies op, waarin de validiteit van 50 klinische testen werd beschreven voor vier aandoeningen:

instabiliteit, intra-articulaire pathology, impingement en rotator cuff scheuren. Er werden geen gegevens gevonden voor andere aandoeningen en de validiteit van items uit de anamnese. De methodologische kwaliteit van de studies, gemeten met de QUADAS, was redelijk. Toch is enige terughoudendheid op zijn plaats, de resultaten waren gebaseerd op sterk geselecteerde patiënten, de selectiecriteria en de uitvoering van de referentie test waren niet of onvoldoende beschreven en de blindering van de beoordelaar was vaak onduidelijk.

In hoofdstuk 5 worden de testresultaten voor instabiliteit en labrum letsels beschreven. Aan de hand van likelihoodratio's (LR) werden de testen met elkaar vergeleken. Voor het diagnosticeren van instabiliteit lijken de *relocation test* en de *anterior release test* waardevol. Minder waardevol zijn de *apprehension test, clunk test, release test, load and shift tests*, en *sulcus sign*. Labrum scheuren lijken het best vast te stellen met de *biceps load I* en de *biceps load II* tests, de *pijn provocatie test van Mimori* en de *endorotatie weerstand test van Zaslav*.

Hoofdstuk 6 bevat de gegevens over validiteit van klinische testen voor het vaststellen van impingement en rotator cuff scheuren. De waarde van deze testen wordt beoordeel aan de hand van de sensitiviteit, specificiteit en achterafkans gebaseerd op Bayes theorema. Een test met een sensitiviteit en specificiteit hoger dan 0.80 en de achteraf kans 0.30 hoger dan de vooraf kans werd als waardevol beoordeeld om als zelfstandig diagnostische instrument te worden gebruikt. Geen van de 15 testen voor het vaststellen van impingement voldeed hieraan. Drie van de 19 testen voor rotator cuff scheuren voldeden aan deze criteria. Dit waren de *Hornblower's sign* en *Dropping sign* voor het vaststellen het niet-operabele scheuren van de M.teres minor of M.infraspinatus en *krachtsverlies in endorotatie* voor het vaststellen van een gedeeltelijke of volledige scheur van de M.infra- of M.supraspinatus.

In hoofdstuk 7 wordt ingegaan op de gevonden resultaten in dit proefschrift. Een steeds terugkerend punt in dit proefschrift was het identificeren van mensen met schouderklachten. Het verschil in definitie

van schouderklachten hinderde een algemene conclusie over de prevalentie van schouderklachten in de review over incidentie en prevalentie van schouderklachten in de open bevolking. De definitie van nieuwe en recidiverende gevallen bepaalde sterk de uitkomst in de cohort studie over het beloop van schouderklachten. Het vaststellen van specifieke aandoeningen aan de schouder werd bemoeilijkt door het ontbreken van een accuraat identificatie systeem, waardoor de behandeling van deze klachten waarschijnlijk niet optimaal zal zijn. Deze punten onderstrepen de behoefte aan consensus over de definitie van schouderklachten, de definitie van een episode voor de beschrijving van het beloop en een classificatiesysteem wat recht doet aan de dagelijkse medische praktijk. Om te komen tegemoet te komen aan deze behoeften zijn een aantal suggesties geformuleerd.

Dankwoord

Dankwoord

Het schrijven van dit proefschrift was een inspirende leerschool. Enerzijds doordat ik mijn eerste systematische review uitvoerde, anderzijds door het schrijven van mijn eerste artikel.
Voor het resultaat ben ik velen dank verschuldigd en niet in het minst aan de onderzoekers die mij voorgingen. Ik heb dankbaar gebruik gemaakt van hun kennis op het gebied van schouderklachten en epidemiologie. Speciaal ben ik dank verschuldigd aan de onderzoekers en mensen die direct betrokken waren bij het schrijven van dit proefschrift.

Bart, dank voor de vrijheid die me hebt gegeven in het uitvoeren van dit project. Je liet me steeds de ruimte om te ontdekken wat de mogelijkheden en beperkingen waren. Je stelde de juiste vragen om me te laten nadenken over de keuzes die ik moest maken.

Lex, dank voor de kennis die je overdroeg in dit project. Je daagde me uit, liet me over ‘muren’ heen kijken en was er altijd om me te helpen wanneer dat nodig was (en dat doe je nog steeds). Het is prettig met je te werken.

Jan, als promotor was je dicht bij dit onderzoek betrokken. In hoeveelheid tijd minder dan Bart en Lex, maar in kwaliteit zeer gewaardeerd. Dank voor deze betrokkenheid en voor de mogelijkheden die je creëerde om contact te leggen met de orthopedie afdeling in het ziekenhuis.

Arianne, in 1996 was je net aan je promotie onderzoek begonnen toe je mij als student gezondheidswetenschappen onder je hoedde kreeg. Vijf jaar later vonden we elkaar terug, hier in Rotterdam. Fijn dat je in mijn projectteam zat. Dank voor je werk bij het beoordelen van artikelen. Je ervaring zorgde er voor dat we mild bleven.

Ingrid, mijn eerste kamergenote en dagelijks begeleider. Samen beleefden we een leuke eerste tijd bij het AKB, helaas vertrok je alweer snel. Dank voor je steun, misschien kruizen onze wegen in de toekomst opnieuw.

Judith, mijn tweede kamergenoot en dagelijks begeleider. Je moest instappen op een rijdende trein (die je ook dagelijks vervoert tussen Amsterdam en Rotterdam). Dat was best lastig, maar je hebt je plaats gevonden. Ik ben blij met de manier waarop je me begeleidde en de ruimte die je me gaf om mijn weg te vinden. Als kamergenoten waren we een mooi stel. Hoogte- en dieptepunten wisselden elkaar af, maar gelukkig zorgde Sacha en Isabella altijd voor een vrolijke noot. Ik ben trots dat je je kwaliteiten als fotograaf beschikbaar stelde voor dit proefschrift. Ik hoop je werk nog vaak te zien.

Harald, als directeur en projectlid was je nauw betrokken bij dit project. Je zorgde dat ik binnen de grenzen van mijn project bleef. Dit resulteerde in 2 ½ jaar in dit boekje. Dank voor je steun en het project bij het Decennium van het Bewegingsapparaat.

Penny, it was a great opportunity to use the QUADAS in its one before last version. The use of this tool made our study stronger. Thanks a lot.

David, thank you for the enormous amount of work you did in editing and upgrading of our manuscript. It took us quite a while to come to the final version, but the manuscript changed to a nice and easy readable article. It was a pleasure to work with you, and to see how you turned dry sentences into lively texts.

Lauraine en David, dank voor jullie hulp bij het corrigeren van het engels. Zonder jullie had ik het niet gered.

Collega's, jullie zorgden voor een gezellige en leuke sfeer binnen het AKB - Decennium. De koffie pauzes waren een welkome onderbreking van de dag en tijdens lunchtijd werd er stevig gewandeld. De positieve invloed van Ingrid is gebleven. Hoewel, Els kon weleens roet in het eten strooien, door op een bankje te willen zitten.
We groeiden uit van een klein clubje tot een echte vrouwengroep. Gelukkig werd dat na het vertrek van Peter gecompenseerd door Sander, Pieter, John en Frank van P&O. Mannen, dank voor de broodnodige afleiding.
Hoewel mijn vertrek ook binnen het AKB niet zo werd gewaardeerd, zit er ook een voordeel aan (elk nadeel heb z'n voordeel). Jullie zijn verlost van het zingen van liedjes op je verjaardag.

Peter en Elske, we waren 1+1+1 = 5. De meest uiteenlopende conversaties voerden we tijdens de lunchpauzes. Ze werden gekenmerkt door een hoog 'van-de-hak-op-de-tak-gehalte' en rare associaties, die altijd voor nieuwe associaties zorgden. Het is prettig me door jullie gesteund te voelen tijdens de promotie plechtigheid.

Peet, dank voor je onuitputtelijke hoeveelheid relativerende woorden. Je wist altijd een positieve kant te belichten als ik alleen de negatieve zag. Ik hoop dat je me nog vaak laat lachen.

Elske, ik zal je missen als mijn onderzoeksmaatje. Ik heb erg genoten van onze wetenschappelijke en minder wetenschappelijke conversaties. Ik hoop je te zien uitgroeien tot een wetenschapper van formaat.

Jos, dankjewel dat je er bent. Ik leerde je kennen als mijn klimmaatje. Samen ontdekten we de balans die nodig was om de top van 'de muur' te bereiken. Je hebt me tijdens het schrijven van dit proefschrift geroped als een klimmaatje. Ik hoop dat we nog lang balanceren.

‘Weet je’ zei de wandelende tak tegen de eekhoorn in de top van de beukenboom, ‘als je helemaal alleen bent kun je niet goed nadenken.’ De eekhoorn keek hem vragend aan en wist niet goed wat hij zeggen moest.
‘Ja’, zei hij toen, een beetje aarzelend.
‘Nou ja, ik bedoel’, ging de wandelende tak verder, ‘dat je dan steeds hetzelfde denkt’.

Toon Tellegen

Curriculum Vitae

Jolanda Luime werd geboren op 6 december 1972 te Ouddorp. In 1990 startte ze haar opleiding fysiotherapie. Ze vervolgde haar studie in 1994 aan de Universiteit Maastricht, waar ze gezondheidswetenschappen studeerde. Deze ronde ze in 1996 af met een pilotstudie naar het effect van bekkenbanden bij patiënten met Spondylitis Ankylopoetica (ziekte van Bechterew). De eerste jaren daarna werkte ze als fysiotherapeut in diverse praktijken in Zuid-Holland en Zeeland. Midden 1998 startte ze bij Fysergo als bedrijfsfysiotherapeut en midden 1999 werd ze ergonomisch adviseur. Ze adviseerde bedrijven bij nieuwbouw, verbouwing en aanpassing van werkplekken. In 2001 besloot ze meer aandacht te willen besteden aan wetenschappelijke kennis over het ontstaan en blijven bestaan van klachten aan het bewegingsapparaat. Ze werd aangesteld als wetenschappelijk medewerker bij het Kenniscentrum voor Arbeid Klachten aan het Bewegingsapparaat op het project ‘evidence based handelen bij nek- en schouderklachten’. Ze deed onderzoek naar de prevalentie, incidentie, beloop en diagnostiek bij schouderklachten. Dit resulteerde in 2004 in dit proefschrift ‘Shoulder complaints. The occurrence, course and diagnosis’. Daarna was ze korte tijd werkzaam voor het Decennium van het Bewegingsapparaat, op dit moment werkt ze als onderzoeker aan de University of Warwick in Groot-Britannië.